

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:03:25 ; Search time 3016.51 Seconds
(without alignments)
17018.784 Million cell updates/sec

Title: US-09-981-353-47

Perfect score: 1764

Sequence: 1 ctgagcctgacggtccgag.....tgaaaaaaaggg 1764

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_ats.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1764	100.0	1764	6	AR113044	BC113044 Sequence
2	1743.2	98.8	1768	9	BC018999	BC018999 Homo sapi
3	1740.2	98.7	1760	9	AK000184	AK000184 Homo sapi
4	1723	97.7	1746	6	AX281695	AX281695 Sequence
5	1360.4	77.1	1362	6	AX301999	AX301999 Sequence
6	993.4	56.3	1758	6	AX305693	AX305693 Sequence
7	856.6	48.6	863	10	MMASML3A	Y08135 M.musculus
8	856.6	48.6	863	6	AX329813	AX329813 Sequence
9	856.6	48.6	863	6	AX330318	AX330318 Sequence
10	856.6	48.6	863	9	HSASML3A	Y08136 H.sapiens m
11	816.4	46.3	1095	6	AX302000	AX302000 Sequence
12	628.2	35.6	181342	2	AC023556	AC023556 Homo sapi
13	553.2	31.4	559	6	AX385445	AX385445 Sequence
14	536.8	30.4	540	6	AX386014	AX386014 Sequence
15	252.4	14.3	359	6	AX260953	AX260953 Sequence
16	243.8	13.8	1310	6	AX247505	AX247505 Sequence
17	243.8	13.8	1610	6	AX332610	AX332610 Sequence
18	243.8	13.8	1610	6	AX335855	AX335855 Sequence
19	243.8	13.8	1610	9	HSASML3B	Y08134 H.sapiens m
20	242.6	13.8	1489	6	AX247501	AX247501 Sequence
21	235.4	13.3	1948	10	BC009087	BC009087 Mus muscu
22	218.2	12.4	122974	2	AC025114	AC025114 Homo sapi
23	218.2	12.4	214308	2	AL732434	AL732434 Homo sapi
24	210	11.9	1566	9	BC014444	BC014444 Homo sapi
25	187.2	10.6	181342	2	AC023556	AC023556 Homo sapi
26	147	8.3	2350	9	AK096144	AK096144 Homo sapi
27	144.4	8.2	156	6	AX261402	AX261402 Sequence
28	131.8	7.5	214308	2	AL732434	AL732434 Homo sapi
29	103.2	5.9	145430	2	AL845421	AL845421 Danio rer
30	103.2	5.9	192318	2	AL845339	AL845339 Plasmodi
31	73.8	4.2	86827	3	PFMAL3P5	AL034556 Plasmodi
32	70.8	4.0	139111	9	AL158048	AL158048 Human DNA
33	70.6	4.0	107289	2	AC116923	AC116923 Dictyoste
34	69.2	3.9	171187	2	AC116960	AC116960 Dictyoste
35	67.8	3.8	10213	3	PFU27338	AE001425 Plasmodi
36	66.4	3.8	11829	3	AE001425	AE001425 Plasmodi
37	66	3.7	110000	2	PFMAL13P2_2	Continuation (3 of
38	65.2	3.7	110000	2	PFMAL13P2_1	Continuation (2 of
39	65	3.7	113117	9	AL512288	AL512288 Human DNA
40	64.8	3.7	14867	3	AE001398	AE001398 Plasmodi
41	64.8	3.7	92633	2	PFMAL4P1_3	Continuation (4 of
42	64.8	3.7	106434	3	AC117080	AC117080 Dictyoste
43	64.4	3.7	14551	6	AX281473	AX281473 Sequence
44	64	3.6	29016	2	AC117266	AC117266 Dictyoste
45	63.8	3.6	92376	8	AC006282	AC006282 Arabidops

ALIGNMENTS

RESULT 1
AR113044
LOCUS AR113044 1764 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 13 from patent US 6132964.
ACCESSION AR113044
VERSION AR113044.1 GI:14093366
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1764)
AUTHORS Bandman,O., Lal,P., Hillman,J.L., Corley,N.C., Guegler,K.J. and Shan,P.
TITLE Hydrolase enzymes
JOURNAL Patent: US 6132964-A 13 17-OCT-2000;

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nslri.nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tjongon, E.E., Touchman, J.W., Teurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 30 Row: a Column: 7
 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, Similarity but not identity to protein.

FEATURES

Location/Qualifiers
 1..1768
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:20681 IMAGE:3138813"
 /tissue_type="Placenta, choriocarcinoma"
 /clone_lib="NIH MGC_21"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"

CDS

78..1439
 /codon_start=1
 /product="Unknown (protein for MGC:20681)"
 /protein_id="AAH1899.1"
 /db_xref="GI:17512053"
 /translation="MALVRALVCLLTAMHCRSLGLFVAPAGRNPPPPAIGQFHWHT
 DLHDPYHTTDDTKYKVCSSKGANANPGFDVLCDSFYQILSAFDFIKNSGGEA
 SFMTGDSPPHVPPELSTDVINTMTNTTQSLFNPLOVPFALGNHDPQOL
 PVTSKYVANLWKPWDEALSTLRKGGYSQVTTNPRLIISLNLVYGNL
 MTNKTDPANQFWELESTLANSQNKKEKYLIAHPVGYLPSSQNTATREYNEKLI
 DIFQKSDVLAGQFYGTHRDSIMVLSKRGSPVNSLFVAPVTPVKSVLKQTNPPG
 IRLFYDPRDYKLDMLQYILNLTANLKGESIKWLEYILTYDIEDLQPESLYGLA
 KQFTILDSKOFIKYNYFFVSYDSSVTCDKTCKAFQICAIMNLNLSIYADCLKQLYIK
 HNY"

BASE COUNT 540 a 358 c 346 g 524 t

ORIGIN

Query Match 98.8%; Score 1743.2; DB 9; Length 1768;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1756; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 1 CTCAGGCTGACGGTCCGAGTGGAGCTGCGGGACAGCCGAACTCCAGTCCAGCCCGC 60
 10 CTCAGGCTGACGGTCCGAGTGGAGCTGCGGGACAGCCGAACTCCAGTCCAGCCCGC 69
 61 GGCCCTCCATGGCGCTGGTGGCGGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 70 GGCCCTCCATGGCGCTGGTGGCGGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 129
 121 GCTCCGCGCTGGCGCTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 130 GCTCCGCGCTGGCGCTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 189
 181 GACAGTTTGGCATGTGACTGACTTACATTTAGACCCCTACTTACACATCAGATGACC 240
 190 GACAGTTTGGCATGTGACTGACTTACATTTAGACCCCTACTTACACATCAGATGACC 249

241 ACACAAAGTGTGTGCTTCATCTAAAGTGCATAATCCCTCCAAACCTCGGCCCTTTGGAG 300
 250 ACACAAAGTGTGTGCTTCATCTAAAGTGCATAATCCCTCCAAACCTCGGCCCTTTGGAG 309
 301 ATGTTCTGTGTGATCTCCATATCAACTTATTTTGTGACGATTTGATTTTAAATTT 360
 310 ATGTTCTGTGTGATCTCCATATCAACTTATTTTGTGACGATTTGATTTTAAATTT 369
 361 CTGGACAAGAAGCATCTTTTCATGATATGACAGGGGATAGCCACCTCATGTTCTGTAC 420
 370 CTGGACAAGAAGCATCTTTTCATGATATGACAGGGGATAGCCACCTCATGTTCTGTAC 429
 421 CTGAACCTCTCAACAGACACTGTTTATAAATGTGATCACTAATATGACAAACCATCCAGA 480
 430 CTGAACCTCTCAACAGACACTGTTTATAAATGTGATCACTAATATGACAAACCATCCAGA 489
 481 GTCTCTTTCCAAATCTCCAGGTTTCCCTCGCGTGGGTAAATCATGACTATTGGCCACAGG 540
 490 GTCTCTTTCCAAATCTCCAGGTTTCCCTCGCGTGGGTAAATCATGACTATTGGCCACAGG 549
 541 ATCAACTGCTGTAGTCACAGTAAAGTGTACAAATGCAGTAGCAAACTCTGGAACCAT 600
 550 ATCAACTGCTGTAGTCACAGTAAAGTGTACAAATGCAGTAGCAAACTCTGGAACCAT 609
 601 GGCTAGATGAAGAAGCTATTAGTACTTTAAGGAAAGTGGTTTTTATTACAGAAAGTTA 660
 610 GGCTAGATGAAGAAGCTATTAGTACTTTAAGGAAAGTGGTTTTTATTACAGAAAGTTA 669
 661 CAACTAATCCAACTTAGGATCATCAGTCTTAAACACAACTTGTACTACGGGCCAAATA 720
 670 CAACTAATCCAACTTAGGATCATCAGTCTTAAACACAACTTGTACTACGGGCCAAATA 729
 721 TAATGACACTGACACAGACTGACCCAGCCCAACAGTTTGAATGGCTAGAAAGTACATTGA 780
 730 TAATGACACTGACACAGACTGACCCAGCCCAACAGTTTGAATGGCTAGAAAGTACATTGA 789
 781 ACAACTCTCAGCAGATAAAGGAGAGGTGTATATCATAGCACATGTTTCCAGTGGGGTATC 840
 790 ACAACTCTCAGCAGATAAAGGAGAGGTGTATATCATAGCACATGTTTCCAGTGGGGTATC 849
 841 TGGCATCTTTCACAGAACTCAGACCAATGAGAGATCTATTAATGAGAAATTTGATAGATA 900
 850 TGGCATCTTTCACAGAACTCAGACCAATGAGAGATCTATTAATGAGAAATTTGATAGATA 909
 901 TTTTTCAAAAATACAGTGTGCTATTGACAGGCAATTTTATGACACACATCTACAGAGACA 960
 910 TTTTTCAAAAATACAGTGTGCTATTGACAGGCAATTTTATGACACACATCTACAGAGACA 969
 961 GCATTATGTTCTTTTCAGATAAAAAAGGAAGTCCAGTAAATTTCTTTTGTGGCTCCTG 1020
 970 GCATTATGTTCTTTTCAGATAAAAAAGGAAGTCCAGTAAATTTCTTTTGTGGCTCCTG 1029
 1021 CTGTTACACAGTGAAGAGTGTTTTGAAGAAAAACAGACCAACCTCTCGTATCAGACTGT 1080
 1030 CTGTTACACAGTGAAGAGTGTTTTGAAGAAAAACAGACCAACCTCTCGTATCAGACTGT 1089
 1081 TTCAGTATGATCCTCGTGATTTAAATTTTGGATATGTTGAGTATTACTTGAATCTGA 1140
 1090 TTCAGTATGATCCTCGTGATTTAAATTTTGGATATGTTGAGTATTACTTGAATCTGA 1149
 1141 CAGAGGCGAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCTTGCACCAACCT 1200
 1150 CAGAGGCGAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCTTGCACCAACCT 1209
 1201 ACACATTGAAGATTTCAGCCGGAAGTTTATATGGAATTAGCTAAACAAATTTACAATCC 1260
 1210 ACACATTGAAGATTTCAGCCGGAAGTTTATATGGAATTAGCTAAACAAATTTACAATCC 1269
 1261 TAGACAGTAGCAGGTTTATAAATACTACAAATTTCTTTGTGAGTTATGACAGCAGTG 1320
 1270 TAGACAGTAGCAGGTTTATAAATACTACAAATTTCTTTGTGAGTTATGACAGCAGTG 1329

REMARK COMMENT

QY 961 GCATTATGTTCTTTTCAGATAAAAGGAAGTCCAGTAAATTTCTTTGTTGGTCTCTG 1020
Db 964 GCATTATGTTCTTTTCAGATAAAAGGAAGTCCAGTAAATTTCTTTGTTGGTCTCTG 1023
QY 1021 CTGTTACACGAGTGAAGAGTGTGTTAGAAAACAGACAAATCTCTGGTATCAGACTGT 1080
Db 1024 CTGTTACACGAGTGAAGAGTGTGTTAGAAAACAGACAAATCTCTGGTATCAGACTGT 1083
QY 1081 TTCAGTATGATCTCTCGTGAATATAATTTATTGGATATGTTGCAGTATCTGTAATCTGA 1140
Db 1084 TTCAGTATGATCTCTCGTGAATATAATTTATTGGATATGTTGCAGTATCTGTAATCTGA 1143
QY 1141 CAGAGGCGAATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGACCT 1200
Db 1144 CAGAGGCGAATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGACCT 1203
QY 1201 ACACATTTGAAGATTTTCAGCCGCAAGATTTATATGATGATGATGATGATGATGATGATG 1260
Db 1204 ACACATTTGAAGATTTTCAGCCGCAAGATTTATATGATGATGATGATGATGATGATGATG 1263
QY 1261 TAGACAGTAAAGCAGTTTATAAAATCTACAAATTTCTTTTGTGAGTTATGACAGCAGTG 1320
Db 1264 TAGACAGTAAAGCAGTTTATAAAATCTACAAATTTCTTTTGTGAGTTATGACAGCAGTG 1323
QY 1321 TAACATGTGATAGACATGTAAGCCCTTTGAGATTTGCAATTTATGAAATCTTTGATAATA 1380
Db 1324 TAACATGTGATAGACATGTAAGCCCTTTGAGATTTGCAATTTATGAAATCTTTGATAATA 1383
QY 1381 TTTCTTATGAGATTTGCTCAACAGCTTTTATATAAGCAATTTACTAGTATTTTACAG 1440
Db 1384 TTTCTTATGAGATTTGCTCAACAGCTTTTATATAAGCAATTTACTAGTATTTTACAG 1443
QY 1441 TTTTGTGCTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTGTTGGAATTTTACATA 1500
Db 1444 TTTTGTGCTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTGTTGGAATTTTACATA 1503
QY 1501 AATCTTTGTTAATTAAGTGGGCAAGTGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1559
Db 1504 AATCTTTGTTAATTAAGTGGGCAAGTGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1563
QY 1560 TTTCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1619
Db 1564 TTTCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1623
QY 1620 GTGCTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
Db 1624 GTGCTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1683
QY 1680 ATCTAATTTGTACCTTTGTTGAATTTGTTGAATTTGTTGAATTTGTTGAATTTGTTGAAT 1739
Db 1684 ATCTAATTTGTACCTTTGTTGAATTTGTTGAATTTGTTGAATTTGTTGAATTTGTTGAAT 1743
QY 1740 AATATGAAAAA 1756
Db 1744 AAAAAA 1760

RESULT 4
AX281695 1746 bp DNA linear PAT 02-NOV-2001
LOCUS Sequence 104 from Patent WO0177389.
DEFINITION AX281695
ACCESSION AX281695
VERSION AX281695.1 GI:16608946
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shiffman, D., Somogyi, R., Lawn, R., Seilhamer, J.J., Porter, G.J.,
Mikita, T. and Tai, J.
TITLE Genes expressed in foam cell differentiation

JOURNAL Patent: WO 0177389-A 104 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1..1746
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 233778-9"
BASE COUNT 521 a 356 c 342 g 527 t
ORIGIN
Query Match 97.7%; Score 1723; DB 6; Length 1746;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1745; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 1 CTCAGGCTGACGGTCCGAGTGGAGCTGCGGACAGCCCGAACCTCCAGGTGAGCCCCG 60
Db 1 CTCAGGCTGACGGTCCGAGTGGAGCTGCGGACAGCCCGAACCTCCAGGTGAGCCCCG 60
QY 61 GGCCCTCCATGCGCTGGTGGCGGCAC-TCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 119
Db 61 GGCCCTCCATGCGCTGGTGGCGGCAC-TCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 120 CGCTCCGGCTCGGGCTGCGGCTGCGGCGCGCAGCGCAGGAATCTCTCCGCGGATA 179
Db 121 CGCTCCGGCTCGGGCTGCGGCTGCGGCGCGCAGCGCAGGAATCTCTCCGCGGATA 180
QY 180 GGACAGTCTTGGCATGTGACTGACTTACCTTAGACCCCTACTTACCACATCAGATGAC 239
Db 181 GGACAGTCTTGGCATGTGACTGACTTACCTTAGACCCCTACTTACCACATCAGATGAC 240
QY 240 CACACAAAGTGTGCTTCTCACTAAAGTGCCTCAACCTGCGGCTGCGGCTGCGGCTGCGG 299
Db 241 CACACAAAGTGTGCTTCTCACTAAAGTGCCTCAACCTGCGGCTGCGGCTGCGGCTGCGG 300
QY 300 GATGTTCTGTGATTTCTCCATATCACTTATTTGTCAGCATTTGATTTTATTAATAAT 359
Db 301 GATGTTCTGTGATTTCTCCATATCACTTATTTGTCAGCATTTGATTTTATTAATAAT 360
QY 360 TCTGACAAAGATCTTTTCATGATGACAGGGATAGCCACCTCATGTTCTCTGTA 419
Db 361 TCTGACAAAGATCTTTTCATGATGACAGGGATAGCCACCTCATGTTCTCTGTA 420
QY 420 CCTGAATCTCAACAGACACTGTTTATAATGTGATCACTAATGACAAACACCATCCAG 479
Db 421 CCTGAATCTCAACAGACACTGTTTATAATGTGATCACTAATGACAAACACCATCCAG 480
QY 480 AGTCTCTTTCAAAATCTCCAGGTTTTCCTGCGCTGGGTAATCATGACTATTGCGCCACAG 539
Db 481 AGTCTCTTTCAAAATCTCCAGGTTTTCCTGCGCTGGGTAATCATGACTATTGCGCCACAG 540
QY 540 GATCAACTGCTGTAGTCAACAGTAAAGTGTACATGAGTAAAGTAAAGTAAAGTAAAGTAAAG 599
Db 541 GATCAACTGCTGTAGTCAACAGTAAAGTGTACATGAGTAAAGTAAAGTAAAGTAAAGTAAAG 600
QY 600 TGGCTAGATGAAGAGCTATTAGTACTTTAAGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGT 659
Db 601 TGGCTAGATGAAGAGCTATTAGTACTTTAAGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGT 660
QY 660 ACACTAATCAAACTTAGGATCATGCTCTAAACAACTTGTACTACGCGCCCAAT 719
Db 661 ACACTAATCAAACTTAGGATCATGCTCTAAACAACTTGTACTACGCGCCCAAT 720
QY 720 ATATGACACTGAACAGAGTCAACAGTAAAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 779
Db 721 ATATGACACTGAACAGAGTCAACAGTAAAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780
QY 780 AACAACTCTCAGCAGAAATAGGAGAGGTGTATATCATAGCACAATGTTCCAGTGGGGTAT 839
Db 781 AACAACTCTCAGCAGAAATAGGAGAGGTGTATATCATAGCACAATGTTCCAGTGGGGTAT 840
QY 840 CTGCCATCTTCAGACAACTCAGCAATGAGAAATCTATATAGAGAAATTTGATAGAT 899
Db 841 CTGCCATCTTCAGACAACTCAGCAATGAGAAATCTATATAGAGAAATTTGATAGAT 900


```
Db      841  |||||AAATACAGTGATGTCATTTGCAGGACAAATTTATGACACACTCACAGAGACGATTTATG 900
Qy      969  GTTCTTTTTCAGATAAAGGAGTCCAGTAAATCTTTTGTGTTGGCTCTGCTGTGTACA 1028
Db      901  GTTCTTTTTCAGATAAAGGAGTCCAGTAAATCTTTTGTGTTGGCTCTGCTGTGTACA 960
Qy     1029  CCAGTGAAGAGTGTGTTTATAGAAAAACAGACCAACAATCCTGGTATCAGACTGTTTCAGTAT 1088
Db     961  CCAGTGAAGAGTGTGTTTATAGAAAAACAGACCAACAATCCTGGTATCAGACTGTTTCAGTAT 1020
Qy     1089  GATCCTCGTGAATTAATAATTTATGGATATGTTGCAGTATTTACTTGAATCTGCAGAGGGG 1148
Db     1021  GATCCTCGTGAATTAATAATTTATGGATATGTTGCAGTATTTACTTGAATCTGCAGAGGGG 1080
Qy     1149  AATCTAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGAGCTTACGACATT 1208
Db     1081  AATCTAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGAGCTTACGACATT 1140
Qy     1209  GAAGATTGCGAGCGGGAAGTTTATATGGAATAGCTAAACAATTTACAATCTCTAGACAGT 1268
Db     1141  GAAGATTGCGAGCGGGAAGTTTATATGGAATAGCTAAACAATTTACAATCTCTAGACAGT 1200
Qy     1269  AAGCAGTTTATAAATACTACAATTACTTCTTGTGAGTTATGACAGCAGTGTAAACATGT 1328
Db     1201  AAGCAGTTTATAAATACTACAATTACTTCTTGTGAGTTATGACAGCAGTGTAAACATGT 1260
Qy     1329  GATAAGACATGTAAAGGCTTTTTCAGATTGTGCAATTAATGAATCTTGATAATTTTCCTAT 1388
Db     1261  GATNAGACATGTAAAGGCTTTTTCAGATTGTGCAATTAATGAATCTTGATAATTTTCCTAT 1320
Qy     1389  GCAGATTGCTCAACAGCTTTTATATAAAGCACAATTTACTAG 1430
Db     1321  GCAGATTGCTCAACAGCTTTTATATAAAGCACAATTTACTAG 1362

RESULT 6
AX305693
LOCUS      AX305693      1758 bp      DNA      linear      PAT 11-DEC-2001
DEFINITION Sequence 444 from Patent WO0188188.
ACCESSION AX305693
VERSION    AX305693.1 GI:17645124
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1
AUTHORS   Iehikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE     Method for examining ischemic conditions
JOURNAL   Patent: WO 0188188-A 444 22-NOV-2001;
           School Juridical Person Nihon University (JP)
FEATURES   Location/Qualifiers
            source
              1..1758
              /organism="Mus musculus"
              /db_xref="taxon:10090"
BASE COUNT 493 a 415 c 386 g 462 t 2 others
ORIGIN

Query Match 56.3%; Score 993.4; DB 6; Length 1758;
Best Local Similarity 75.6%; Pred. No. 3.1e-159;
Matches 1334; Conservative 1; Mismatches 393; Indels 36; Gaps 7;

Qy      1  CTCAGGCTTGACGGTCCGAGTGGAGCTGCGGACAGCCGAACTCCAGGTCAGCCCCG 60
Db      27  CTAGGGCTTGGGGTCTCAGCGGCGCCCAAGCTCCAGTCTAGCTGCTCTGG-GGCAGTGG 84
Qy     61  GGCCCTCCATGGCGCTGTGGCGGCACCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCC 120
Db     85  CCTCCACCATGGCGTGTGCTGGCAACTTCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 144
Qy    121  GCTCCGGCTGCGGCTGCGGCGCGGCGGAGAAATTCCTTCCTCCGCGGATAG 180
```

```
Db      145  GCGCGGGGCTCGGGGTGCCCCCTGGCGCGCGC-----GGATCGCGCCCCCAGCTGTGG 195
Qy     181  GACAGTTTTTGGCATGTGACTGACTTACACTTACAGCCCTACTTACACATCACAGATGACC 240
Db     196  GGCAGTTTTTGGCAGCGTACTGACTTACATCTAGACCCCTACTTACACATTTACAGATGACC 255
Qy     241  ACACAAAAGTGTGTGCTTTCATCTAAAGGTGCAAAATGCTCCAAACCTCGGCCCTTTTGGAG 300
Db     256  GTACCAAGGTGTGTGCTTTCATCTAAAGGTGCAAAATGCTCCAAACCTCGGCCCTTTTGGAG 315
Qy     301  ATGTCTGTGTGATCTCTCCATATCACTTTATTTTGTGCAGATTTGATTTTATTAATAAT 360
Db     316  ATGTCTGTGTGATCTCTCCATATCACTTTATTTTGTGCAGCTTTGATTTTATTAAGAAAT 375
Qy     361  CTGGACAAGAAGCATCTTTTCATGATATGACAGGGGATAGCCACCTCATGTTTCTGTATC 420
Db     376  CAGGACAAGAAGCATCTTTTCATGATATGACAGGGGATAGCCACCTCATGTTGCGCAGTAC 435
Qy     421  CTGAATCTCTCAACAGACACTGTTTATAAATGTGATCACTAATATGACAAACCATCCAGA 480
Db     436  CTGAATCTCTCAACAGACCGTGTATAAAGGTGATCACTAATATGACAAATGACTGTCCAGA 495
Qy     481  GTCTCTTTCCAAATCTCCAGGTTTCCCTCGCGTGGTAAATCATGACTATTGGCCACAGG 540
Db     496  ACCTGTGTTCCAAACCTCCAGGTTTCCCTCGCACTGGGCAATCATGACTATTGGCCACAGG 555
Qy     541  ATCAACTGCTGCTAGTACACAGTAAAGTGTACAAATGCAGTAGCAAAACCTCTGGAACCAT 600
Db     556  ACCAGTGGCAATAGTACACAGTAAAGTGTACAGTGTCTGCTGCTGCTGCTGCTGCTGCTG 615
Qy     601  GGCTAGATGAAGAAGCTATTAGTACTTTTAAAGGAAGGTGGTGTGTTTATTCACAGAAAGTTA 660
Db     616  GGCTGGGTGAAGAAGCTATTAGTACTTTTAAAGGAAGGTGGTGTGTTTACTCACAGAAAGTTG 675
Qy     661  CAACTAATCCAAACCTTAGATCATCAGTCTAAACACAACTTGTTACTAGGCCCAATA 720
Db     676  CAACTAATCCAGCTTTGAGGATCATTTAGCTTAAACACAACTTGTTACTAGGCCCAATA 735
Qy     721  TAATGACACTGAACAAGAGTGAACCCAGCCCAACAGTTTCAATGGCTAGAAAGTACATTTGA 780
Db     736  TCATGACCTTGAACAAGACAGACCCAGCAANTCAGTTTGAATGGCTGGAAATACACTCA 795
Qy     781  ACAACTCTCAGCAGATAAAGGAGAGGTGTATATCATATAGCACATGTTTCCAGTGGGTATC 840
Db     796  ACAGCTCTCTATGGAATAAGGAGAGGTATACWTATCATGCGCATGTTTCCAGTGGGTATC 855
Qy     841  TGCCATCTTCACAGAACATCACAGCAATGAGAGATCTATATATGAGAAATTTGATAGATA 900
Db     856  TCCTTTATGCAACTGACACCCCGCGGATTAAGGCAGTACTATATATGAGAAACTGCTTTGATA 915
Qy     901  TTTTTCAAAAATACAGTGATGTGCTTTCAGGAGCAATTTTATGACACACTTCACAGAGACA 960
Db     916  TTTTTCAGAGATACAGCTCGGTCATTTGCGGACAGTTCTATGGCCACACCCATAGAGACA 975
Qy     961  GCATTTATGGTTCCTTCAGATAAAAAAGGAAGTCCAGTAAATTTCTTTGTTTGGTCTCTG 1020
Db     976  GCCTTATGTCTCTTCAGATAAAGCGGAATCCCACTCAATTTCTGTGTTTGTGGCACCTG 1035
Qy    1021  CTGTTACACAGTGAAGAGTGTGTTTAAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db    1036  CTGTTACACAGTGAAGAGGAGTGTGTTTAAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095
Qy    1081  TTCAAGTATGATCTCTGCTGATTAATAATTTATGATATTTGAGTATTTACTTCAATCTGA 1140
Db    1096  TTCAAGTATGATCTCTGCTGATTAATAATTTATGATATTTGAGTATTTACTTCAATCTGA 1155
Qy    1141  CAGAGGCGAATCTAAAGGAGAGTCCATCTCGAAGCTGGAGTATATCTTCCAGCCAGACCT 1200
Db    1156  CAGAGGCGAATCTAAAGGAGAGTCCATCTCGAAGCTGGAGTATATCTTCACTCAGGCT 1215
Qy    1201  ACAGATTTGAAGATTTGACGCGGAGGAGTATATGATATTTAGTAAACAAATTTTCAATCC 1260
Db    1216  ACAGTGTTCAGATCTGCAGCCCAAGAGTTTATATGCTTCTAGTTTCAAGCAATTTGCCACCA 1275
```



```
Db 856 TCCCTATGCAACTGACACCCCGGCGATAGGCGAGTACTATAATGAGAACTGCTTGATA 915
Qy 901 TTTTTCAAAAATACAGTGATGTCATTGCGAGGACAAATTTATGGACACACTCACAGACA 960
Db 916 TTTTCAGAAGATACAGCTCCGTCATTCGCGGACAGTCTATGGCCACACCCCATAGAGACA 975
Qy 961 GCATTATGGTCTTTTCAGATAAAAAAGGAAGTCCAGTAAATCTTTTGTGTGGCTCTG 1020
Db 976 GCCTTATGGTCTTTTCAGATAAAGCGGAATCCACTCAATCTGTGTGTGGCACTG 1035
Qy 1021 CTGTTACACCAAGTGAAGAGTGTTTTAGAAAAACAGACCAAACTCTGGTATCAGACTGT 1080
Db 1036 CTGTTACACCAAGTGAAGAGTGTTTACAAAAGGAGACCAAACTCCGGTGTCCGCTAT 1095
Qy 1081 TTCAGTATGATCTCTGCTGATTATATAATTTATGGATATGTTGCAGTATTTACTTGAATCTGA 1140
Db 1096 TTCAGTACAAGCTGGTGATTACATATGCTGGACATGGTGCAATATTTACTTGAACCTGA 1155
Qy 1141 CAGAGGCGAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGACCT 1200
Db 1156 CAGAAGCCAAATCTAAAGGAGAACTCAACTGGACATTTGGAGTATGCTTGACTCAGGCT 1215
Qy 1201 ACAGCATTGAAGATTTGCGAGCGGAAAGTTTATATGATAGCTTAAACAATTTACAATCC 1260
Db 1216 ACAGTGTGAGATCTGCGAGCCAAAGAGTTTATATGCTTGTAGTTTCAGCAATTTGCCACCA 1275
Qy 1261 TAGACAGTAAAGCAGTTTATAAAATPACTACAATTTACTTCTTTGTGAGTTATGACAGAGTG 1320
Db 1276 AAGACAGCAAGCAGTCTCTGAAATPACTACCATTTATTTTGTGAGTTATGACAGAGTG 1335
Qy 1321 TACATCTGATACAGATGTAAGGCTTTTCAAGTTTGTGCAATTTATGAATTTTGAATAA 1380
Db 1336 CAATCTGTGACAGCATTGTAAGACCTTTACAGGCTCTGTGCAATTTATGAATCTTTGATAGA 1395
Qy 1381 TTTCTATGAGATTTGCTCAACAGCTTTTATATAAGCACAATTTACTAGTATTTACAG 1440
Db 1396 TGTCTATGATGATTTGCTTTAAACAGATT-TATAAGCAGCTCACTAGTATTTCCAGTT 1454
Qy 1441 TTTTGTCTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTGTGGAAATTTTACATA 1500
Db 1455 GTGTTTGTAGAAAAA-----AAATCACATCATGGTCTCTCGAT 1497
Qy 1501 AATCTTTGTTAAATTTACTAGTGGGCAAGTAGACTTCTGTCTTTGCTTTCTTTTCTTTT 1560
Db 1498 CAGC-TTCTGAGTGTGGGAGCATAGGAAGATTGACATTTCTGTGTGACTTTGTGAGGCTC 1556
Qy 1561 TCTTTTGTATGATGATGATATCTTTATCATTTCTGAAATTTGATTTATATATTTAAAG 1620
Db 1557 AGATGCTGATACC---AGTAACATTCAAAATTTAACTTCACGCATTTGAATACATTTAAAC 1612
Qy 1621 TGCTCATTATAGAAATGATGATGTAATTTGGATGTAATATTC--AGTTTATATAATTA 1678
Db 1613 TGTTCATTAGCGGAGTATATTTGGACATAAAATATCTGCTTTCCAATTTATGTAATTA 1672
Qy 1679 TATCTAATTTGTACCTCTGTGAAATTTGCAATTTATACAAATAAGGAAATCTTTATCTC 1738
Db 1673 TACCTAACTTATACCTCTGTGAAACTTTGATCTATGCAATTAAGCAGTTGATTTTCTG 1732
Qy 1739 TAAATATGAAAAAATAAAAAA 1762
Db 1733 TGAATAAAAAAATAAAAAA 1756
```

```
RESULT 8
AX329813
LOCUS .
DEFINITION Sequence 322 from Patent WO0194629.
ACCESSION AX329813
VERSION AX329813.1 GI:18102791
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
```

```
REFERENCE
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 322 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source 1..863
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 285 a 125 c 148 g 305 t
ORIGIN
```

```
Query Match 48.6%; Score 856.6; DB 6; Length 863;
Best Local Similarity 99.5%; Pred. No. 2.1e-170;
Matches 859; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 895 TAGATATTTTTCAAAAATACAGTATGTCATTGCGAGGACAAATTTATGGACACACTCACA 954
Db 1 TAGATATTTTTCAAAAATACAGTATGTCATTGCGAGGACAAATTTATGGACACACTCACA 60
Qy 955 GAGACAGCATTATGTTCTTTTCAGATAAAAAAGGAAGTCCAGTAAATTTCTTTGTTGG 1014
Db 61 GAGACAGCATTATGTTCTTTTCAGATAAAAAAGGAAGTCCAGTAAATTTCTTTGTTGG 120
Qy 1015 CTCCTGCTGTTTACACAGTGAAGAGTGTTTTAGAAAAACAGACCAACAACTCTGATCA 1074
Db 121 CTCCTGCTGTTTACACAGTGAAGAGTGTTTTAGAAAAACAGACCAACAACTCTGATCA 180
Qy 1075 GACTGTTTCAGTATGATCTCTGATATATAAATTTATGATATGTTGAGTATTTACTTGA 1134
Db 181 GACTGTTTCAGTATGATCTCTGATATATAAATTTATGATATGTTGAGTATTTACTTGA 240
Qy 1135 ATCTGACAGAGCGCAATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCC 1194
Db 241 ATCTGACAGAGCGCAATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCC 300
Qy 1195 AGACCTACGACATTGAAGATTTCAGCGGAAAGTTTATATGATGATGATGATGATGATGAT 1254
Db 301 AGACCTACGACATTGAAGATTTCAGCGGAAAGTTTATATGATGATGATGATGATGATGAT 360
Qy 1255 CAATCTGACAGTGAAGCAGTTTATAAATTAATACTACAATTTCTTTGTGAGTATGACA 1314
Db 361 CAATCTGACAGTGAAGCAGTTTATAAATACTACAATTTCTTTGTGAGTATGACA 420
Qy 1315 GCAGTGAACATGTCATAAGACATGTAAGGCTTTTCAGATTTCGCAATTATGATGATGAT 1374
Db 421 GCAGTGAACATGTCATAAGACATGTAAGGCTTTTCAGATTTCGCAATTATGATGATGAT 480
Qy 1375 ATATATTTTCTATGACAGATTGCTCAACAGCTTTTATATAAAGCACAATTTACTAGTAT 1434
Db 481 ATATATTTTCTATGACAGATTGCTCAACAGCTTTTATATAAAGCACAATTTACTAGTAT 540
Qy 1435 TCACAGTTTTTGTCTAATAGAAATGCTGATCTGATTTCTGAGATCAATTTGTGGGAATTT 1494
Db 541 TCACAGTTTTTGTCTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTGTGGGAATTT 600
Qy 1495 TACATAAATCTTTGTTAAATTTACTGAGTGGGCAAGTAGACTTCTGCTGCTTTGCTTTCTTT 1554
Db 601 TACATAAATCTTTGTTAAATTTACTGAGTGGGCAAGTAGACTTCTGCTGCTTTGCTTTCTTT 660
Qy 1555 TTTTCTTTTCTGATGCTTTAATGATGATATCTTTTATCAATCTGAAATTTGATTTATATAT 1614
Db 661 TTTTCTTTTCTGATGCTTTAATGATGATATCTTTTATCAATCTGAAATTTGATTTATATAT 720
Qy 1615 TTAAGTGTCTCAATTAATAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1674
Db 721 TTAAGTGTCTCAATTAATAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 1675 ATTATATCTAATTTGTACCTTTGTTGAAATTTGTTTATATATATAAAGCAATTTCTTTA 1734
```

```

Db      781 ATTATATCTAAATTGTCACCTTGTGCAATTGTCATTATACATAAAGGAATTCCTTTA 840
QY      1735 TCTCTAAATATGAAAAA 1757
Db      841 TCTCTAAAAA 863

RESULT 9
LOCUS   AX330318                863 bp    DNA        linear    PAT 09-JAN-2002
DEFINITION Sequence 827 from Patent WO0194629.
ACCESSION AX330318
VERSION   AX330318.1 GI:18103296
KEYWORDS  human.
SOURCE   Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS  Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
          Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE    Cancer gene determination and therapeutic screening using signature
          gene sets
JOURNAL  Patent: WO 0194629-A 827 13-DEC-2001;
          Avalon Pharmaceuticals (US)
FEATURES
source   Location/Qualifiers
          1..863
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
BASE COUNT 285 a 125 c 148 g 305 t
ORIGIN

Query Match 48.6%; Score 856.6; DB 6; Length 863;
Best Local Similarity 99.5%; Pred. No. 2.1e-170;
Matches 859; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      895 TAGATATTTTCAAAATACAGTGATCTCATTCGAGCAAAATTTTATGACACACTCACA 954
Db      1 TAGATATTTTCAAAATACAGTGATCTCATTCGAGCAAAATTTTATGACACACTCACA 60
QY      955 GAGACAGCATTTATGGTCTTTTCAGATAAAAAAGAGTCCAGTAAATCTCTTTGTTGG 1014
Db      61 GAGACAGCATTTATGGTCTTTTCAGATAAAAAAGAGTCCAGTAAATCTCTTTGTTGG 120
QY      1015 CTCCTGCTGTACACAGTGAAGAGTGTTTTGAAGAAAACAGACCAATCTCTGGTATCA 1074
Db      121 CTCCTGCTGTACACAGTGAAGAGTGTTTTGAAGAAAACAGACCAATCTCTGGTATCA 180
QY      1075 GACTGTTTCAGTATGATCCTCGGATTATAAATTATTTGGATATGTTGCAGTATTACTTGA 1134
Db      181 GACTGTTTCAGTATGATCCTCGGATTATAAATTATTTGGATATGTTGCAGTATTACTTGA 240
QY      1135 ATCTGACAGAGGCGAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCCTGACCC 1194
Db      241 ATCTGACAGAGGCGAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCCTGACCC 300
QY      1195 AGACCTACGACATTGAAGATTTCAGCCGGAAGTTTATATGATTAGCTTAAACAATTTA 1254
Db      301 AGACCTACGACATTGAAGATTTCAGCCGGAAGTTTATATGATTAGCTTAAACAATTTA 360
QY      1255 CAATCTCAGACAGTAAAGAGTTTATAAATACTACAAATCTCTTTGTTGAGTTATGACA 1314
Db      361 CAATCTCAGACAGTAAAGAGTTTATAAATACTACAAATCTCTTTGTTGAGTTATGACA 420
QY      1315 GCAGTGTAAATGTGATAAGACATGTAAGGCCCTTTTCAGATTGTTGCAATTTATGAATCTTG 1374
Db      421 GCAGTGTAAATGTGATAAGACATGTAAGGCCCTTTTCAGATTGTTGCAATTTATGAATCTTG 480
QY      1375 ATAAATATTTCTATGAGATTGCTCAACAGCTTTATATAAGACCAATTAAGTACTAGTATT 1434
Db      481 ATAAATATTTCTATGAGATTGCTCAACAGCTTTATATAAGACCAATTAAGTACTAGTATT 540

```

```

QY      1435 TCACAGTTTTTCTAATAGAAAATGCTGATTCGTGATTCGTGAGATCAAAATTTGTGGAAATTT 1494
Db      541 TCACAGTTTTTCTAATAGAAAATGCTGATTCGTGATTCGTGAGATCAAAATTTGTGGAAATTT 600
QY      1495 TACATAAAATCTTTGTTTAATTAATCTAGTGGGCAAGTAGACTTCTCTGCTTTGCTTTCTTTT 1554
Db      601 TACATAAAATCTTTGTTTAATTAATCTAGTGGGCAAGTAGACTTCTCTGCTTTGCTTTCTTTT 660
QY      1555 TTTTCTTTCTTTTGATGCTTAATCTAGATATCTTTTATCATCTCTGAATGTATTATATAT 1614
Db      661 TTTTCTTTCTTTTGATGCTTAATCTAGATATCTTTTATCATCTCTGAATGTATTATATAT 720
QY      1615 TTAAGAGTCTCAATTAATAGAAATGATGGATGTAATTTGGATGTAATATTCAGTTTATATA 1674
Db      721 TTAAGAGTCTCAATTAATAGAAATGATGGATGTAATTTGGATGTAATATTCAGTTTATATA 780
QY      1675 ATTATATCTAAATTTGACCTTTGTTGAAATTCATTTATATACAATAAAGCGAATTCCTTTA 1734
Db      781 ATTATATCTAAATTTGACCTTTGTTGAAATTCATTTATATACAATAAAGCGAATTCCTTTA 840
QY      1735 TCTCTAAATATGAAAAA 1757
Db      841 TCTCTAAAAA 863

RESULT 10
LOCUS   HSASML3A                863 bp    mRNA        linear    PRI 19-SEP-1996
DEFINITION H.sapiens mRNA for ASM-like phosphodiesterase 3a.
ACCESSION Y08136
VERSION   Y08136.1 GI:1552272
KEYWORDS  acid sphingomyelinase; ASML3 gene; phosphodiesterase.
SOURCE   Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS  Hofmann, K.
TITLE    Acid Sphingomyelinase is a member of a multi-gene family and shares
          motifs with a large family of metallo-phosphoesterases
JOURNAL  Unpublished
AUTHORS  Hofmann, K.
TITLE    Direct Submission
JOURNAL  Submitted (17-SEP-1996) K. Hofmann, Iserc (Swiss Inst. F. Exp.
          Canc. Res.), Bioinformatics Group, Chemin Des Boveresses 155,
          Ch/1066 Epalinges S/Lausanne, SWITZERLAND
FEATURES
source   Location/Qualifiers
          1..863
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          1..536
          /gene="ASML3a"
          <1..536
          /gene="ASML3a"
          /note="putative"
          /codon_start=3
          /product="acid sphingomyelinase-like phosphodiesterase"
          /protein_id="CAA69330.1"
          /db_xref="GI:1552273"
          /db_xref="SPTREMBL:Q92484"
          /translation="DIFOKYSDVIAGQFVGHTRHDSIMVLSDKGSPVNSLFPVAPAVT
          PVKSVLEKQTNNGIRLFQYDPRDYLKLDMLQYLNLTANLKGSIWKLEVLTOTY
          DIEDQPSLEGLAKQFTILDSKQFIKYNFYFVSDSVTCDKTCAKFQICAINMLD
          NISYADCLKQLYIKHY"
BASE COUNT 285 a 125 c 148 g 305 t
ORIGIN

Query Match 48.6%; Score 856.6; DB 9; Length 863;
Best Local Similarity 99.5%; Pred. No. 2.1e-170;
Matches 859; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      895 TAGATATTTTCAAAATAACAGTGATGTCATTGTCAGGACAAATTTTATGACACACTCACA 954

```


JOURNAL

COMMENT

969 GTTCTTTTCAGATATAAAAAAGGAAGTCCAGTAAATCTTTCTTTGTGTGGCTCCTGCTGTAC 1028
Db 634 GTTCTTTTCAGATATAAAAAAGGAAGTCCAGTAAATCTTTCTTTGTGTGGCTCCTGCTGTAC 693
Qy 1029 CCAGTGAAGAGTGTTTTATAGAAAAACAGACCAACAATCCCTGGTATCAGATCTGTTTCAGTAT 1088
Db 694 CCAGTGAAGAGTGTTTTATAGAAAAACAGACCAACAATCCCTGGTATCAGATCTGTTTCAGTAT 753
Qy 1089 GATCCTCGTATATAAATATTGGATATGTTGCAGTATTACTTGAATCTGCAGAGGCG 1148
Db 754 GATCCTCGTATATAAATATTGGATATGTTGCAGTATTACTTGAATCTGCAGAGGCG 813
Qy 1149 AATCTAAAGGAGTCCATCTGGAAGCTGGAGTATATCCTGACCCAGACCTACGACATT 1208
Db 814 AATCTAAAGGAGTCCATCTGGAAGCTGGAGTATATCCTGACCCAGACCTACGACATT 873
Qy 1209 GAAGATTTTCAGCGCGAAAGTTTATATGATTAGCTAAACAATTTACAACTCCTAGACAGT 1268
Db 874 GAAGATTTTCAGCGCGAAAGTTTATATGATTAGCTAAACAATTTACAACTCCTAGACAGT 933
Qy 1269 AAGCAGTTTATAAATACATAAATTTACTTCTTTGTGAGTTATGACGAGTGTAAATCT 1328
Db 934 AAGCAGTTTATAAATACATAAATTTACTTCTTTGTGAGTTATGACGAGTGTAAATCT 993
Qy 1329 GATAGACATGTAAGGCCCTTCAGATTTCTGCAATTTATGAATCTTGATAATATTCCTAT 1388
Db 994 GATAGACATGTAAGGCCCTTCAGATTTCTGCAATTTATGAATCTTGATAATATTCCTAT 1053
Qy 1389 GCAGATTTGCTCAACAGCTTTATATAAAGCAATTAAGTACTAG 1430
Db 1054 GCAGATTTGCTCAACAGCTTTATATAAAGCAATTAAGTACTAG 1095

RESULT 12

AC023556/c

LOCUS

DEFINITION Homo sapiens chromosome 7 clone RP11-655J4 map 7, *** SEQUENCING IN PROGRESS ***, 54 unordered pieces.

ACCESSION

AC023556

VERSION

AC023556.2

KEYWORDS

HTG; HTGS; PHASE1.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 181342)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 7, clone RP11-655J4

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 181342)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P., DeAvellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M., Fenster,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Larcocque,K., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivari,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Firtell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE

Direct Submission

Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 14, 2000 this sequence version replaced gi:6978252.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L6522

Center clone name: 655_J_4

* NOTE: This is a 'working draft' sequence. It currently
* consists of 54 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1083: contig of 1083 bp in length
* 1084 1183: gap of 100 bp
* 1184 2308: contig of 1125 bp in length
* 2309 2408: gap of 100 bp
* 2409 3852: contig of 1444 bp in length
* 3853 3952: gap of 100 bp
* 3953 5046: contig of 1094 bp in length
* 5047 5146: gap of 100 bp
* 5147 6612: contig of 1466 bp in length
* 6613 6712: gap of 100 bp
* 6713 8248: contig of 1536 bp in length
* 8249 8348: gap of 100 bp
* 8349 9395: contig of 1047 bp in length
* 9396 9495: gap of 100 bp
* 9496 10850: contig of 1355 bp in length
* 10851 10950: gap of 100 bp
* 10951 12677: contig of 1727 bp in length
* 12678 12777: gap of 100 bp
* 12778 14111: contig of 1334 bp in length
* 14112 14211: gap of 100 bp
* 14212 15612: contig of 1401 bp in length
* 15613 15712: gap of 100 bp
* 15713 17225: contig of 1513 bp in length
* 17226 17325: gap of 100 bp
* 17326 17964: contig of 639 bp in length
* 17965 18064: gap of 100 bp
* 18065 19509: contig of 1445 bp in length
* 19510 19609: gap of 100 bp
* 19610 21080: contig of 1471 bp in length
* 21081 21180: gap of 100 bp
* 21181 22551: contig of 1371 bp in length
* 22552 22651: gap of 100 bp
* 22652 24300: contig of 1649 bp in length
* 24301 24400: gap of 100 bp
* 24401 26046: contig of 1646 bp in length
* 26047 26146: gap of 100 bp
* 26147 27868: contig of 1722 bp in length
* 27869 27968: gap of 100 bp
* 27969 29612: contig of 1644 bp in length
* 29613 29712: gap of 100 bp
* 29713 32434: contig of 2722 bp in length
* 32435 32534: gap of 100 bp
* 32535 34326: contig of 1792 bp in length
* 34327 34426: gap of 100 bp
* 34427 36340: contig of 1914 bp in length
* 36341 36440: gap of 100 bp
* 36441 38516: contig of 2076 bp in length
* 38517 38616: gap of 100 bp
* 38617 40173: contig of 1557 bp in length
* 40174 40273: gap of 100 bp

FEATURES

```

location/Qualifiers
1. .181342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-655J4"
/clone_lib="RP11 Human Male BAC"
1. .1083
misc_feature
1184. .2308
/note="assembly_fragment"
misc_feature
2409. .3852
/note="assembly_fragment"
misc_feature
3953. .5046
/note="assembly_fragment"

```

RESULT 13	DEFINITION
AX385445	ACCESSION
LOCUS	VERSION

LOCUS	AX385445	559 bp	DNA	linear	PAT 19-MAR-2002
DEFINITION	Sequence 373 from Patent WO0214500.				
ACCESSION	AX385445				
VERSION	AX385445.1	GI:19578575			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:03:00 ; Search time 255.01 seconds
(without alignments)
15577.944 Million cell updates/sec

Title: US-09-981-353-47
Perfect score: 1764
Sequence: 1 ccaggctgacgtccgag.....Tgaaaaaagagagagagag 1764

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1764	100.0	1764	22	AAC60227
2	1735.8	98.4	1783	24	AAS62765
3	1723	97.7	1746	24	AAS94849
4	1567.4	88.9	1874	22	AAI59324
5	1546.4	87.7	1873	22	AAI61110
6	1360.4	77.1	1362	22	AAD21343
7	993.4	56.3	1758	24	ABI99482
8	927	52.6	939	22	AAH20423
9	856.6	48.6	863	24	ABL84743

10	856.6	48.6	863	24	ABL61985	Colon adenocarcino
11	856.6	48.6	863	24	ABL62490	Colon adenocarcino
12	816.4	46.3	1095	22	AAD21344	Human nuclear fact
13	701.8	39.8	728	21	AAA02374	Human colon cancer
14	657.4	37.3	733	21	AAA02333	Human colon cancer
15	574.6	32.6	778	21	AAA02346	Human colon cancer
16	553.2	31.4	559	24	ABN60406	Human cancer relat
17	536.8	30.4	540	24	ABN60375	Human cancer relat
18	534.6	30.3	773	21	AAA02332	Human colon cancer
19	300	17.0	300	21	AAA00750	Human colon cancer
20	252.4	14.3	359	23	AAS57928	cDNA #604 encoding
21	243.8	13.8	1610	24	ABL64782	Lung cancer relat
22	243.8	13.8	1610	24	ABL68027	Ovary cancer relat
23	242.6	13.8	1489	22	AAH99916	Nucleotide sequenc
24	242.6	13.8	1814	22	AAI59259	Human polynucleoti
25	242.6	13.8	1816	22	AAI61045	Human polynucleoti
26	144.4	8.2	156	22	AAS58377	cDNA #1053 encodin
27	84.8	4.8	756	21	AAA02416	Human colon cancer
28	67.8	3.8	9789	17	AAI41852	cDNA encoding Plas
29	64.4	3.7	14551	24	ABL34584	Human metastasis a
30	62.4	3.5	6286	22	AAS46591	Tumour suppressor
31	62.4	3.5	13377	22	AAS46476	Tumour suppressor
32	62.4	3.5	13377	24	ABL33463	Human immune syste
33	62	3.5	15674	24	ABL70513	Chemically treated
34	62	3.5	15674	24	ABL32362	Human immune syste
35	62	3.5	15674	24	ABL34476	Human metastasis a
36	61.8	3.5	340	23	ABV13635	Human prostate exp
37	61.4	3.5	556	23	ABV40063	Human prostate exp
38	61.4	3.5	556	23	ABV40163	Human prostate exp
39	61.4	3.5	556	23	ABV42105	Human prostate exp
40	61.4	3.5	556	23	ABV43601	Human prostate exp
41	61.4	3.5	7544	22	AAS45301	Chemically pretrea
42	61.4	3.5	7544	24	ABK28140	DNA transcription
43	61.2	3.5	10328	24	ABL33544	Human immune syste
44	61	3.5	1902	23	ABL13819	Drosophila melanog
45	60.8	3.4	7351	24	ABL32028	Human immune syste

ALIGNMENTS

RESULT 1
AAC60227
ID AAC60227 standard; cDNA; 1764 BP.

AC AAC60227;

XX 13-FEB-2001 (first entry)

DE Human hydrolase-like molecule 5 cDNA.

XX Hydrolase-like molecule; human; cell proliferation disorder;
KW autoimmune; cancer; AIDS; acquired immune deficiency syndrome; ss.

XX Homo sapiens.

OS US6132964-A.

XX 17-OCT-2000.

PF 06-FEB-1998; 98US-0013881.

XX 06-FEB-1998; 98US-0013881.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Hillman JL, Guegler KJ, Shah P, Lal P, Corley NC;
DR WPI; 2001-006133/01.

XX P-PSDB; AAB28794.

XX New human hydrolase-like molecules (HHLMs) and polynucleotides encoding
PT the HHLMs, useful for diagnosing, treating or preventing cell

DT	14-FEB-2002 (first entry)	
XX	cdna sequence #552 encoding novel human secreted protein.	
DE	Human secreted protein; hyperproliferative disorder; autoimmune disorder;	
XX	immune deficiency disorder; blood disorder; inflammatory disorder;	
KW	infectious disorder; gene therapy; antimicrobial; hepatotropic;	
KW	immunosuppressive; antirheumatic; ss.	
XX	Homo sapiens.	
OS		
XX	WO200177291-A2.	
PN	18-OCT-2001.	
XX		
XX	29-MAR-2001; 2001WO-US10485.	
PF	06-APR-2000; 2000US-195604P.	
XX	(GEMY) GENETICS INST INC.	
PA		
XX	Wong GG, Clark HF, Fecthel K, Agostino MJ, Howes SH, Resnick RJ;	
PI	Gullukota K, Graham JR;	
PI	WPI; 2002-010900/01.	
XX		
DR	New polynucleotides encoding secreted proteins useful for treating e.g.	
PT	asthma, HIV and Crohn's disease -	
PT	Claim 1; Page 356; 391pp; English.	
XX		
CC	The present invention relates to the isolation of novel cDNA sequences	
CC	which encode human secreted proteins. The cDNA sequences have been	
CC	derived from a variety of human tissues. The invention also provides	
CC	a method for producing proteins from these polynucleotide sequences.	
CC	The proteins are useful for identifying compounds that modulate their	
CC	activity and production, and the cell is also useful for identifying	
CC	compounds that modulate expression of the polynucleotide sequences	
CC	encoding the secreted proteins. The sequences of the invention are	
CC	useful for treating diseases such as hyperproliferative disorders	
CC	(e.g. cancer), immune deficiency disorders (e.g. severe combined	
CC	immunodeficiency (SCID)), autoimmune disorders (e.g. multiple	
CC	sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory	
CC	disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).	
CC	The polynucleotide sequences of the invention are also useful in gene	
CC	therapy. AAS62214-AAS62838 represent the cDNA sequences of the	
CC	invention that encode for novel human secreted proteins.	
XX		
SQ	Sequence 1783 BP; 523 A; 376 C; 348 G; 536 T; 0 other;	
	Query Match 98.4%; Score 1735.8; DB 24; Length 1783;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 1737; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 CTCAGGCGCTGACCGTCCGAGTGAGCTGGGACAGCCGAACTCCAGGTGAGCCCGGC 60	
DB	45 CTCAGGCGCTGACCGTCCGAGTGAGCTGGGACAGCCGAACTCCAGGTGAGCCCGGC 104	
QY	61 GGCCCTCCATGGCGCTGGTGGCGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 120	
DB	105 GGCCCTCCATGGCGCTGGTGGCGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 164	
QY	121 GCTCCGGCTCGGGCTGGCGCTGGCGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 180	
DB	165 GCTCCGGCTCGGGCTGGCGCTGGCGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 224	
QY	181 GACAGTTTGGCATGTGACTGACTTACACTTAGACCCCTACTTACACATCAGATGACC 240	
DB	225 GACAGTTTGGCATGTGACTGACTTACACTTAGACCCCTACTTACACATCAGATGACC 284	
QY	241 ACACAAAGTGTGCTTCATCTAAAGGTGCAATGCTCCAACTGGCCCTTTTGGAG 300	
DB	285 ACACAAAGTGTGCTTCATCTAAAGGTGCAATGCTCCAACTGGCCCTTTTGGAG 344	
QY	301 ATGTTCTGTGTGATTCTCCATATCACTTATTTTGTGACGATTTGATTTTATTTAAATTT 360	
DB	345 ATGTTCTGTGTGATTCTCCATATCACTTATTTTGTGACGATTTGATTTTATTTAAATTT 404	
QY	361 CTGGACAAGAAGCATCTTTTCATGATATGGACAGGGATAGCCACCTCATGTTCTCTGTAC 420	
DB	405 CTGGACAAGAAGCATCTTTTCATGATATGGACAGGGATAGCCACCTCATGTTCTCTGTAC 464	
QY	421 CTGAACCTCTCAACAGACACTGTTTAAATGTGATCACTAATATGACAAACACATCCAGA 480	
DB	465 CTGAACCTCTCAACAGACACTGTTTAAATGTGATCACTAATATGACAAACACATCCAGA 524	
QY	481 GTCTCTTTCCAAATCTCCAGGTTTTCCCTGGCTGGTAAATCATGACATTTGGCCACAGG 540	
DB	525 GTCTCTTTCCAAATCTCCAGGTTTTCCCTGGCTGGTAAATCATGACATTTGGCCACAGG 584	
QY	541 ATCAACTGCTGTAGTCACCAAGTAAAGTGTACAATGCAAGTAGCAAAACCTCTGGAACCAT 600	
DB	585 ATCAACTGCTGTAGTCACCAAGTAAAGTGTACAATGCAAGTAGCAAAACCTCTGGAACCAT 644	
QY	601 GGCTAGATGAAGAAGCTATTAGTACTTTAAAGGAAGGTGGTTTTTATTTACAGAAAGTTA 660	
DB	645 GGCTAGATGAAGAAGCTATTAGTACTTTAAAGGAAGGTGGTTTTTATTTACAGAAAGTTA 704	
QY	661 CAACTAATCCAAACCTTAGGATCATGCTCTAAACACAAACTTGTACTACGGCCCAATA 720	
DB	705 CAACTAATCCAAACCTTAGGATCATGCTCTAAACACAAACTTGTACTACGGCCCAATA 764	
QY	721 TAATGACACTGAACAAGACTGACCCAGCCACCACTTTGAATGGCTAGAAAGTACATTGA 780	
DB	765 TAATGACACTGAACAAGACTGACCCAGCCACCACTTTGAATGGCTAGAAAGTACATTGA 824	
QY	781 ACAACTCTCAGCAGAATAAGGAAGGTGTATATCATAGCACATGTTCCAGTGGGGTATC 840	
DB	825 ACAACTCTCAGCAGAATAAGGAAGGTGTATATCATAGCACATGTTCCAGTGGGGTATC 884	
QY	841 TGCCATCTTACAGAACATCACAGCAATGAGAAATCACTATATAGAGAAATTGATAGATA 900	
DB	885 TGCCATCTTACAGAACATCACAGCAATGAGAAATCACTATATAGAGAAATTGATAGATA 944	
QY	901 TTTTTCAAAATACAGTGTGTCATTTGACGAGCAATTTTATGACACACTCACAGAGACA 960	
DB	945 TTTTTCAAAATACAGTGTGTCATTTGACGAGCAATTTTATGACACACTCACAGAGACA 1004	
QY	961 GCATTATGGTCTTTTCAGATAAAAGGAAGTCCAGTAAATCTTTTGTGGCTCCTG 1020	
DB	1005 GCATTATGGTCTTTTCAGATAAAAGGAAGTCCAGTAAATCTTTTGTGGCTCCTG 1064	
QY	1021 CTGTTACACAGTCAAGAGTGTTTTAAAGGAAGTCCAGTAAATCTTTTGTGGCTCCTG 1080	
DB	1065 CTGTTACACAGTCAAGAGTGTTTTAAAGGAAGTCCAGTAAATCTTTTGTGGCTCCTG 1124	
QY	1081 TTCAGTATGATCTCGTGTATTAATAATTATGGATATGTTGAGTATTAATCTTGAATCTGA 1140	
DB	1125 TTCAGTATGATCTCGTGTATTAATAATTATGGATATGTTGAGTATTAATCTTGAATCTGA 1184	
QY	1141 CAGAGCGCAATCTAAAGGGAGAGTCCATCTGGAGCTGAGTATATCTCTGAGCCAGACCT 1200	
DB	1185 CAGAGCGCAATCTAAAGGGAGAGTCCATCTGGAGCTGAGTATATCTCTGAGCCAGACCT 1244	
QY	1201 ACAGCATTTGAGATTTGACCCGGAAGTTTATGGATTTAGCTAAACAAATTTACAATCC 1260	
DB	1245 ACAGCATTTGAGATTTGACCCGGAAGTTTATGGATTTAGCTAAACAAATTTACAATCC 1304	
QY	1261 TAGACAGTAAAGCAGTTTATAAATACTCAATTAATCTTCTTGTGAGTTATGACAGCAGTG 1320	
DB	1305 TAGACAGTAAAGCAGTTTATAAATACTCAATTAATCTTCTTGTGAGTTATGACAGCAGTG 1364	
QY	1321 TAACATGTGATGAACATGTGAAGGCTTTTCAGATTTGTGCAATTTATGAATCTTGATAATA 1380	
DB	1365 TAACATGTGATGAACATGTGAAGGCTTTTCAGATTTGTGCAATTTATGAATCTTGATAATA 1424	

Db 961 AGCATTATGTTCTTTTTCAGATAAAAAAGGAGTCCAGTAAATCTTTTGTGTGGTCTCT 1020
Qy 1020 GCTGTTACACAGTGAAGAGTGTGTTTATAGAAAAACAGAACCAATCTGCTGATCAGACTG 1079
Db 1021 GCTGTTACACAGTGAAGAGTGTGTTTATAGAAAAACAGAACCAATCTGCTGATCAGACTG 1080
Qy 1080 TTTTCAGTATGATCTCTGATGATTAATAATTTATGATATGTTGTCAGTATTAATCTGAAATCTG 1139
Db 1081 TTTTCAGTATGATCTCTGATGATTAATAATTTATGATATGTTGTCAGTATTAATCTGAAATCTG 1140
Qy 1140 ACAGAGCGGAATCTTAAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTTGCAGCCAGACC 1199
Db 1141 ACAGAGCGGAATCTTAAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTTGCAGCCAGACC 1200
Qy 1200 TAGCACATTTGAAGATTTGCGAGCCGGAAGTTTATATGATTTAGTCAATTAACAATCTTCAATC 1259
Db 1201 TAGCACATTTGAAGATTTGCGAGCCGGAAGTTTATATGATTTAGTCAATTAACAATCTTCAATC 1260
Qy 1260 CTAGACAGTGAAGAGTGTGTTTATAAATACTACAAATCTTCTTTGTCAGTATGACAGCAGT 1319
Db 1261 CTAGACAGTGAAGAGTGTGTTTATAAATACTACAAATCTTCTTTGTCAGTATGACAGCAGT 1320
Qy 1320 GTAACATGTCATAGACATGTAAGGCTTTTCAGATTTGTCGAATTAATGATCTTGTATAAT 1379
Db 1321 GTAACATGTCATAGACATGTAAGGCTTTTCAGATTTGTCGAATTAATGATCTTGTATAAT 1380
Qy 1380 ATTTCTCTATGAGATTTGCCTCAACAGCTTTATATAAGCACCAATTAATGATTTTTCACA 1439
Db 1381 ATTTCTCTATGAGATTTGCCTCAACAGCTTTATATAAGCACCAATTAATGATTTTTCACA 1440
Qy 1440 GTTTTCTCTATAGAAAATGCTGATTTCTGATTTCTGAGATCAATTTGCGAAATTTTACAT 1499
Db 1441 GTTTTCTCTATAGAAAATGCTGATTTCTGATTTCTGAGATCAATTTGCGAAATTTTACAT 1500
Qy 1500 AATCTTTGTTAATTAATGAGTGGGCAAGTAGACTTCTGCTGTTGCTTTCTTTTCTTTT 1559
Db 1501 AATCTTTGTTAATTAATGAGTGGGCAAGTAGACTTCTGCTGTTGCTTTCTTTTCTTTT 1559
Qy 1560 TTTCTTTGATGCTTAAATGATATATCTTTATCATTTCTGAATTTGATATATATTTAAA 1619
Db 1560 TTTCTTTGATGCTTAAATGATATATCTTTATCATTTCTGAATTTGATATATATTTAAA 1619
Qy 1620 GTGCTCAATTAATGAGATGAGATGTAATTTGAGATGTAATTTTCAAGTATATATAATAT 1679
Db 1620 GTGCTCAATTAATGAGATGAGATGTAATTTGAGATGTAATTTTCAAGTATATATAATAT 1679
Qy 1680 ATCTAATTTGTACCTTTGTTGAATTTGTCATTTATACAAATAAGCGAATTTCTTTATCTCT 1739
Db 1680 ATCTAATTTGTACCTTTGTTGAATTTGTCATTTATACAAATAAGCGAATTTCTTTATCTCT 1739
Qy 1740 AATATG 1746
Db 1740 AATATG 1746

RESULT 4

AAI59324 standard; cDNA; 1874 BP.

ID AAI59324

XX AAI59324;

AC AAI59324;

XX 22-OCT-2001 (first entry)

DT Human polynucleotide SEQ ID NO 1527.

DE Human, nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

OS Homo sapiens.
XX WO200153312-A1.
PN 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
PF 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM40168.
DR Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PT Claim 1; SEQ ID NO 1527; 10078pp; English.
PS The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as; Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 1874 BP; 596 A; 357 C; 334 G; 587 T; 0 other;
SQ Query Match 88.9%; Score 1567.4; DB 22; Length 1874;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 175 CGATAGGACAGTGTGTCATCTAAAGGTGCAATGCTCCCAACCCCTGGCCCTT 294
Db 298 CAACAGGACAGTGTGTCATCTAAAGGTGCAATGCTCCCAACCCCTGGCCCTT 417
Qy 235 ATGACCACACAAAGTGTGTCATCTAAAGGTGCAATGCTCCCAACCCCTGGCCCTT 294
Db 358 ATGACCACACAAAGTGTGTCATCTAAAGGTGCAATGCTCCCAACCCCTGGCCCTT 417
Qy 295 TTGGAGATGTTCTGTGTGATTTCCATATCAATTTTGTGTCAGATTTGATTTTATTA 354
Db 418 TTGGAGATGTTCTGTGTGATTTCCATATCAATTTTGTGTCAGATTTGATTTTATTA 477
Qy 355 AAAATTTCTGACAAGAAGCATCTTTTCATGATGACAGGGGATGACCCACCTCATGTTTC 414
Db 478 AAAATTTCTGACAAGAAGCATCTTTTCATGATGACAGGGGATGACCCACCTCATGTTTC 537
Qy 415 CTGTACCTGAATCTTCAACAGACACTGTTTAAATGTGTACTTAATATGACACACCA 474
Db 538 CTGTACCTGAATCTTCAACAGACACTGTTTAAATGTGTACTTAATATGACACACCA 597

QY	475	TCAGAGTCTCTTCCAAATCTCCAGGTTTCCCTGCGTGGTAAATCATGACTATTGGC	534	QY	1555	TTTTTTCTTTTGATGCTTAATAGTAGATATCTTTATCATCTCAATTCGATTTATATATAT	1614
Db	598	TCAGAGTCTCTTCCAAATCTCCAGGTTTCCCTGCGTGGTAAATCATGACTATTGGC	657	Db	1678	TTTTTTCTTTTGATGCTTAATAGTAGATATCTTTATCATCTCAATTCGATTTATATATAT	1737
QY	535	CACAGGATCAACTGCGCTAGTACACAGTAAGAGTGATCAATAGCAGTAGCAAACTCTGGA	594	QY	1615	TAAAGTCTCATTAATAGATGATGGATGTAATAATGGATGTAATAATTCAGTTTATATA	1674
Db	658	CACAGGATCAACTGCTGCTAGTACACAGTAAGAGTGATCAATAGCAGTAGCAAACTCTGGA	717	Db	1738	TAAAGTCTCATTAATAGATGATGGATGTAATAATGGATGTAATAATTCAGTTTATATA	1797
QY	595	AACCATGGCTAGATGAAGAGCTATTAGTACTTTTAAGGAAAGTGTTTATTATTCACAGA	654	QY	1675	ATTATATCTAAATTTGTACCCCTGTTGAAATTTGTCATTTATACAATAAAGCGAATTCCTTTA	1734
Db	718	AACCATGGCTAGATGAAGAGCTATTAGTACTTTTAAGGAAAGTGTTTATTATTCACAGA	777	Db	1798	ATTATATCTAAATTTGTACCCCTGTTGAAATTTGTCATTTATACAATAAAGCGAATTCCTTTA	1857
QY	655	AAAGTTACAACTAATCCAAACCTTAGGATCATCAGTCTAAACACAACTTGTACTAGGCGC	714	QY	1735	TCTCTAAATATGAAAAA	1751
Db	778	AAAGTTACAACTAATCCAAACCTTAGGATCATCAGTCTAAACACAACTTGTACTAGGCGC	837	Db	1858	TCTCTAAAAA	1874
QY	715	CAATATATATGACACTGAACAAGACTGACCCAGCAACAGTTTGAATGGCTAGAAAGTA	774	RESULT 5			
Db	838	CAATATATATGACACTGAACAAGACTGACCCAGCAACAGTTTGAATGGCTAGAAAGTA	897	AAI61110			
QY	775	CATTGAACAACTCTCAGCAGATAAGGAGAGGTGTATATCATAGCACATGTTCCAGTGG	834	ID	AAI61110	standard; cDNA; 1873 BP.	
Db	898	CATTGAACAACTCTCAGCAGATAAGGAGAGGTGTATATCATAGCACATGTTCCAGTGG	957	XX	AAI61110;		
QY	835	GGTATCTGCATCTTACAGACATCAGCAGCAATGAGCAATCTATATAGAAATTCGA	894	AC	AAI61110;		
Db	958	GGTATCTGCATCTTACAGACATCAGCAGCAATGAGCAATCTATATAGAAATTCGA	1017	XX	22-OCT-2001	(first entry)	
QY	895	TAGATATTTTTCAAAAATACAGTGATGTCATTCGAGGACAAATTTTATGGACACACTCACA	954	XX	Human polynucleotide SEQ ID NO 5099.		
Db	1018	TAGATATTTTTCAAAAATACAGTGATGTCATTCGAGGACAAATTTTATGGACACACTCACA	1077	DE	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
QY	955	GAGACAGCAATTAATGTTCTTTTTCAGATAAAAAAGGAGTCCAGTAATTTCTTTTGTGG	1014	KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
Db	1078	GAGACAGCAATTAATGTTCTTTTTCAGATAAAAAAGGAGTCCAGTAATTTCTTTTGTGG	1137	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
QY	1015	CTCCTGCTGTACACAGTGAAGAGTGTATTAAGAAAACAGACCAATCTCGGTATCA	1074	KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
Db	1138	CTCCTGCTGTACACAGTGAAGAGTGTATTAAGAAAACAGACCAATCTCGGTATCA	1197	KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
QY	1075	GACTGTTTCAGTATGATCTCTGATTTATAAATTTATGGATGTTTGCAGTATTACTTGA	1134	XX	leukemia; ss.		
Db	1198	GACTGTTTCAGTATGATCTCTGATTTATAAATTTATGGATGTTTGCAGTATTACTTGA	1257	OS	Homo sapiens.		
QY	1135	ATCTGACAGAGGGAATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATCTCTGACCC	1194	XX	WO200153312-A1.		
Db	1258	ATCTGACAGAGGGAATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATCTCTGACCC	1317	PD	26-JUL-2001.		
QY	1195	AGACCTTACGACATTTGAAGATTTGCGAGCGGAAAGTTTATATGATTTAGCTAAACAATTTA	1254	XX	26-DEC-2000; 2000WO-US34263.		
Db	1318	AGACCTTACGACATTTGAAGATTTGCGAGCGGAAAGTTTATATGATTTAGCTAAACAATTTA	1377	XX	21-JAN-2000; 2000US-0488725.		
QY	1255	CAATCTCAGACAGTAAGCAGTTTATAAATACTACAAATTTCTTTTGTGAGTTATGACA	1314	PR	25-APR-2000; 2000US-0552317.		
Db	1378	CAATCTCAGACAGTAAGCAGTTTATAAATACTACAAATTTCTTTTGTGAGTTATGACA	1437	PR	09-JUL-2000; 2000US-0598042.		
QY	1315	CGAGTGTAAATGTGATAAGACATGTAAGCGCTTTCAGATTTTGTGCAATTTATGAATCTTG	1374	PR	19-JUL-2000; 2000US-0620312.		
Db	1438	CGAGTGTAAATGTGATAAGACATGTAAGCGCTTTCAGATTTTGTGCAATTTATGAATCTTG	1497	PR	03-AUG-2000; 2000US-0653450.		
QY	1375	ATAATATTTTCTATGAGATTCGCTCAACAGCTTTATATAAGCACAATTTACTAGTATT	1434	PR	14-SEP-2000; 2000US-0662191.		
Db	1498	ATAATATTTTCTATGAGATTCGCTCAACAGCTTTATATAAGCACAATTTACTAGTATT	1557	PR	19-OCT-2000; 2000US-0693036.		
QY	1435	TCACAGTTTTTGTCTAATGAAAAATGCTGATTTCTGATTTCTGAGATCAATTTGTGGGAATTT	1494	PR	29-NOV-2000; 2000US-0727344.		
Db	1558	TCACAGTTTTTGTCTAATGAAAAATGCTGATTTCTGATTTCTGAGATCAATTTGTGGGAATTT	1617	XX	(HYSE-) HYSEQ INC.		
QY	1495	TACATAAATCTTTTGTAAATCTGAGTGGGCAAGTAGACTTCTGCTTTTGTCTTTT	1554	XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
Db	1618	TACATAAATCTTTTGTAAATCTGAGTGGGCAAGTAGACTTCTGCTTTTGTCTTTT	1677	PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
				PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
				XX	WPI; 2001-442253/47.		
				DR	P-PSDB; AAM41954.		
				XX	Novel nucleic acids and polypeptides, useful for treating disorders		
				PT	such as central nervous system injuries -		
				XX	Claim 1; SEQ ID NO 5099; 10078pp; English.		
				CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and		
				CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,		
				CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
				CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
				CC	of the invention may be used to treat diseases of the peripheral nervous		
				CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
				CC	localised neuropathies and central nervous system diseases, such as		
				CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC C.N.S disorders, arthritis and inflammation, leukaemias and
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX
 SQ Sequence 1873 BP; 596 A; 359 C; 335 G; 583 T; 0 other;

Query Match 87.7%; Score 1546.4; DB 22; Length 1873;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1564; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 175 CGATAGGACAGTTTGGCATGTGACTGACTTACACTTAGACCCCTACTTACACATCACAG 234
 DB 298 CAACAGGACAGTTTGGCATGTGACTGACTTACACTTAGACCCCTACTTACACATCACAG 357
 QY 235 ATGACCACACAAAGTGTGCTTCATCTAAAGTGCATAATGCCCTCCAAACCTGGCCCTT 294
 DB 358 ATGACCACACAAAGTGTGCTTCATCTAAAGTGCATAATGCCCTCCAAACCTGGCCCTT 417
 QY 295 TTGGAGATGTTCTGTGTGATTTCTCCATATCAACTTATTTGTGACGATTTGATTTATTA 354
 DB 418 TTGGAGATGTTCTGTGTGATTTCTCCATATCAACTTATTTGTGACGATTTGATTTATTA 477
 QY 355 AAAATTTCTGGACAAGAGCACTTTTTCATGATATGACAGGGGATAGCCACCTCATGTTTC 414
 DB 478 AAAATTTCTGGACAAGAGCACTTTTTCATGATATGACAGGGGATAGCCACCTCATGTTTC 537
 QY 415 CTGTACCTGAACTCTCAACAGACACTGTTTAAATGTGATCACTAATATGACAACACCA 474
 DB 538 CTGTACCTGAACTCTCAACAGACACTGTTTAAATGTGATCACTAATATGACAACACCA 597
 QY 475 TCCAGAGTCTCTTTTCCAAATCTCCAGGTTTTCCTGCGCTGGTAAATCATGACTATTGGC 534
 DB 598 TCCAGAGTCTCTTTTCCAAATCTCCAGGTTTTCCTGCGCTGGTAAATCATGACTATTGGC 657
 QY 535 CACAGGATCAACTGCTGTAGTACCAAGTAAAGTGTACAAATGACAGTAAACCTCTGGA 594
 DB 658 CACAGGATCAACTGCTGTAGTACCAAGTAAAGTGTACAAATGACAGTAAACCTCTGGA 717
 QY 595 AACCATGGCTAGATGAAGAGCTATTAGTACTTTTAAAGAAAGTGTGTTTATTCACAGA 654
 DB 718 AACCATGGCTAGATGAAGAGCTATTAGTACTTTTAAAGAAAGTGTGTTTATTCACAGA 777
 QY 655 AAGTTACACTAATCAAAACCTTAGGATCATCAGTCTAAACACAACTTTGTACTACGGCC 714
 DB 778 AAGTTACACTAATCAAAACCTTAGGATCATCAGTCTAAACACAACTTTGTACTACGGCC 837
 QY 715 CAAATATATGACACTGAACAGACTGACCCAGCCCAACAGTTTGAATGGCTAGAAAGTA 774
 DB 838 CAAATATATGACACTGAACAGACTGACCCAGCCCAACAGTTTGAATGGCTAGAAAGTA 897
 QY 775 CATTTGAACAACTCTCAGCAGAAATAGGAGAAGTGTATATCATACACATGTTCCAGTGG 834
 DB 898 CATTTGAACAACTCTCAGCAGAAATAGGAGAAGTGTATATCATACACATGTTCCAGTGG 957
 QY 835 GGTATCTGCCATCTTTCACAGAACATCAGCAATGAGAAATGAGAAATGGA 894
 DB 958 GGTATCTGCCATCTTTCACAGAACATCAGCAATGAGAAATGAGAAATGGA 1017
 QY 895 TAGATATTTTTCAAAAATACAGTATGCTTTCAGAGGCAATTTTATGGACACACTCACA 954
 DB 1018 TAGATATTTTTCAAAAATACAGTATGCTTTCAGAGGCAATTTTATGGACACACTCACA 1077
 QY 955 GAGACAGCATTTAGTCTTTTCAGATATAAAAGGAAGTCCAGTAAATCTTTTGTGTTGG 1014
 DB 1078 GAGACAGCATTTAGTCTTTTCAGATATAAAAGGAAGTCCAGTAAATCTTTTGTGTTGG 1137
 QY 1015 CTCCTGCTGTACACAGTGAAGAGTGTTTAGAAAAACAGACCAACCTCCTGGTATCA 1074

Db 1138 CTCCTGCTGTTACACCAGTGAAGAGTGTTTTAGAAAAACAGACCAACAATCCTGGTATCA 1197
 QY 1075 GACTGTTTTCAGTATGATCCTCGTGATATAAAATTTTGGATATGTTGCGAGTATTACTTGA 1134
 DB 1198 GACTGTTTTCAGTATGATCCTCGTGATATAAAATTTTGGATATGTTGCGAGTATTACTTGA 1257
 QY 1135 ATCTGACAGAGGCGAAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAAGTATATCCTGACCC 1194
 DB 1258 ATCTGACAGAGGCGAAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAAGTATATCCTGACCC 1317
 QY 1195 AGACCTTACGACATGGAAGATTTGCGACCGGAAAGTTTATGATGATTTAGTAAACAATTTA 1254
 DB 1318 AGACCTTACGACATGGAAGATTTGCGACCGGAAAGTTTATGATGATTTAGTAAACAATTTA 1377
 QY 1255 CAATCTTAGACAGTAGTAAAGAGTGTATAAAATCTACAAATTTCTTTGAGTATGACA 1314
 DB 1378 CAATCTTAGACAGTAGTAAAGAGTGTATAAAATCTACAAATTTCTTTGAGTATGACA 1437
 QY 1315 GCAGTGTAAATGCTGATAGACATGTAAGGCTTTTCAAGTGTGCAATTTATGAATCTTG 1374
 DB 1438 GCAGTGTAAATGCTGATAGACATGTAAGGCTTTTCAAGTGTGCAATTTATGAATCTTG 1497
 QY 1375 ATAATATTTTCCCTATGACAGATTCGCTCAACAGCTTTTATATAAGCAGCTTTTATATAAGCAGCTTTT 1557
 DB 1498 ATAATATTTTCCCTATGACAGATTCGCTCAACAGCTTTTATATAAGCAGCTTTTATATAAGCAGCTTTT 1557
 QY 1435 TCACAGTGTGCTTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTCTGGGAAATTT 1494
 DB 1558 TCACAGTGTGCTTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTCTGGGAAATTT 1617
 QY 1495 TACATAAATCTTGTGTAATTAAGTCTGAGTGGGCAAGTACAGTTCCTGCTTTCTTTT 1554
 DB 1618 TACATAAATCTTGTGTAATTAAGTCTGAGTGGGCAAGTACAGTTCCTGCTTTCTTTT 1677
 QY 1555 TTTTCTTTCTTTGATGCTTTAATGTAGATATCTTTATCATTTCTGGAATTCGTTATATAT 1614
 DB 1678 TTTTCTTTCTTTGATGCTTTAATGTAGATATCTTTATCATTTCTGGAATTCGTTATATAT 1737
 QY 1615 TTAAGTGTCTTAATAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1674
 DB 1738 TTAAGTGTCTTAATAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1797
 QY 1675 ATTATATCTAATTTGTACCCCTTG-TTGAATTTGTCATTTATACATTAAGCGAATTCCTTT 1733
 DB 1798 ATTATATCTAATTTGTACCCCAAGTTGAACTGTCTATTATATACATTAAGCGAATTCCTTT 1857
 QY 1734 ATCTCTAAATATGAAA 1749
 DB 1858 ATCTCTAAATATGAAA 1873

RESULT 6

AAD21343
 ID AAD21343 standard; cDNA; 1362 BP.

XX AAD21343;

AC AAD21343;

XX AAD21343;

DT 28-JAN-2002 (first entry)

XX Human nuclear factor kappaB-inducing factor (NFIF)-14b cDNA.

DE Human; NFkappaB; nuclear factor kappaB inducing factor; NFIF-14b;

XX NFIF-7a; immune response; inflammatory response; atherosclerosis;

KW rheumatoid arthritis; NSAID-induced gastropathy; scrapie; sepsis;

KW neurodegenerative disease; autoimmune disease; antitense therapy;

KW renal disease; restenosis; brain injury; viral disease; apoptosis;

KW Alzheimer's disease; pleiotropic cytokine; gene therapy; asthma;

KW Crohn's disease; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Key

PH Location/Qualifiers

FT 1..1362

ABI99482

XX ABI99482 standard; cDNA; 1758 BP.

AC ABI99482;

XX DT 07-MAR-2002 (first entry)

XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:444.

XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;

XX KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

XX OS Mus musculus.

XX PN W020018188-A2.

XX PD 22-NOV-2001.

XX PF 18-MAY-2001; 2001WO-JP04192.

XX PR 18-MAY-2000; 2000JP-0145977.

XX PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX DR WPI; 2002-034733/04.

XX DR P-PSDB; ABB57183.

XX PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes -

XX PS Claim 2; Page 1248-1251; 2690pp; English.

XX CC The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI99202 to ABI9912, encoding the protein sequences in ABB57020 to ABB57374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. ABI9913 and ABI9914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention.

XX SQ Sequence 1758 BP; 493 A; 415 C; 386 G; 462 T; 2 other;

Query Match 56.3%; Score 993.4; DB 24; Length 1758;
Best Local Similarity 75.6%; Pred. No. 6.9e-212;
Matches 1334; Conservative 1; Mismatches 393; Indels 36; Gaps 7;

QY 1 CTGAGCGCTGACGGTCCGAGTGGAGCTGCGGCACAGCCGACCTCCAGGTGAGCGCCGCGC 60

DB 27 CTAGGGCGCTCGGGTCTCAGCGCGCGCCAGCTCAGTCCAGTCCGCTCGG--GGCAGTGG 84

QY 61 GGCCTCCATGGCGGTGGTGGCGGCACTCGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

DB 85 CCTCCACCATGGCGGTGGTGGCGGCACTTCCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 144

QY 121 GTCCTGGCTCGGGTGGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180

DB 145 GCGCGGGCTCGGGTGGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 195

QY 181 GACAGTTTGGCATGTGACTGACTTACCTTAGACCTTACTTACCATCATCAGATGACC 240

DB 196 GGCAGTTTGGCAGTGTGACTGACTTACCTTAGACCTTACTTACCATCATCAGATGACC 255

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

241 ACACAAAGTGTGTGCTTCACTAAAGGTGCAATGCTCCAAACCCCTGGCCCTTTTGGAG 300
256 GTACCAAGGTGTGTGCTTCACTAAAGGTGCAATGCTCCAAACCCCTGGCCCTTTTGGAG 315
301 ATGTTCTGTGTGATTTCCATATCAACTTATTTTGTGAGCATTTGATTTTATTAATAATT 360
316 ATGTCCTGTGTGATTTCCATATCAACTTATTTTGTGAGCATTTGATTTTATTAATAATT 375
361 CTGACAAAGAGCATCTTTTCATGATATGACAGGGGATAGCCACCTCATGTTCTGTGTAC 420
376 CAGGACAAAGAGCATCTTTTCATGATATGACAGGGGATAGCCACCTCATGTTCTGTGTAC 435
421 CTGAATCTCAACAGACACTGTTTATAATGTGATCACTAATATGACAAACACCATCCACA 480
436 CTGAATCTCAACAGGACCGTGATAAAGTGTGATCACTAATGATGATGATGATGATGATG 495
481 GTCTCTTTCCAAATCTCCAGGTTTCCCTGCGCTGGTGAATCATGATTTGGCCACAGG 540
496 ACCTGTTTCCAAACCTCCAGGTTTCCCTGCGCTGGGCAATCATGATCTGATGATGATG 555
541 ATCAATGCTGTGTAGTCACCAAGTGTGATAAAGTGTGATAAAGTGTGATAAAGTGTG 600
556 ACCAGCTGCAATAGTCACCAAGTGTGATAAAGTGTGATAAAGTGTGATAAAGTGTG 615
601 GGCTAGATGAAGAGCTATTAGTACTTTAAGAAAGTGTGTTTATTTACAGAAAGTTA 660
616 GGCTGGTGAAGAGCTATTAGTACTTTAAGAAAGTGTGTTTATTTACAGAAAGTTG 675
661 CAACTAATCCAAACCTTAGGATCATGCTCTAAACACAACTGTTACTACGGCCCAATA 720
676 CAACTAATCCAGGCTTGAGGATCATTAGCTTAAACACAACTGTTACTATGGCCCAACA 735
721 TAATGACACTGAACAGACTGACCCAGCCAACTGTTTGAATGGCTAGAAAGTACATTGA 780
736 TCATGACCTGAACAGAGCAGACCCAGCAATCAGTTTGAATGCTGGAATATACACTCA 795
781 ACAACTCTCAGCAGATAAAGGAGAGGTGTATATCATAGACATGTTCCAGTGGGGTATC 840
796 ACAGCTCTATGGAATAAGGAGAGGTGTATACMTATAGCGCATGTTCCAGTGGGGTATC 855
841 TGCCATCTTACAGAAATCAGCAATCAGCAATGAGAGAACTACTAATAGAGAAATGATAG 900
856 TCCCTTATGCAACTGACACCCCGCGGATAGGAGTACTATATGAGAAATGCTGTTGATA 915
901 TTTTTCAAAATACAGTGTGATTCATTCAGAGCAATTTTATGACACACTCACAGAGACA 960
916 TTTTCAGAGATCAGCTCCGTCATTGCGGGAGAGTTCTATGCGCCACACCCCATAGAGA 975
961 GCATTATGGTCTTTTCAGATAAAAAAGGAAAGTCCAGTAAATCTTTGTTGCTGCTCCTG 1020
976 GCCTTATGGTCTTTTCAGATAAANAACGGGATCCACTCAATCTGTTTGTGGCCTCG 1035
1021 CTGTTACACAGTGAAGAGTGTTTTGAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1036 CTGTTACACAGTGAAGAGTGTTTTGAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1095
1081 TCCAGTATGATCTCTGCTGATTAATAATTTATGATATGTTGAGTATTTACTTGAATCTGA 1140
1096 TCCAGTATGATCTCTGCTGATTAATAATTTATGATATGTTGAGTATTTACTTGAATCTGA 1155
1141 CAGAGGGGATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTCTGAGCCAGACCT 1200
1156 CAGAGGGGATCTAAAGGAGAGTCCATCTGGAAGCTGGAGTATGTTGTTGACTCAGGCT 1215
1201 ACAGATTTGAAGATTTGACAGCGGAGAGTATATGAGTATGAGTATGAGTATGAGTATG 1260
1216 ACAGTGTGAGATCTGACAGCGGAGAGTATATGAGTATGAGTATGAGTATGAGTATG 1275
1261 TAGACGTAAAGAGTATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAAT 1320
1276 AAGACAGCAGAGAGTCTCTGAAATTAATAATTAATAATTAATAATTAATAATTAATAAT 1335
1321 TAACATGTGATAGACATGTAAAGGCTTTAGATTTGCAATTAATAATTAATAATTAATAAT 1380

1336 CAACCTTGTCACGACGATTGTAAGACCTTACAGCTCTGTGCAATATGATCTTTGATGCA 1395
1381 TTCTCTATGACGATTGCCTCAACAGCCTTATATAAGACAAATTAAGTATGATTTTACAG 1440
1396 TGCTCTATGATGATTGCTTAAACAGCAAT-TATAAAGCACAGTCTCACTAGTATTCAGTT 1454
1441 TTTTTCCTAATAGAAATCTGATCTGATCTGATCTGATCAATTTGGGCAATTTTACATA 1500
1455 GTGTTGTAGAAAAA-----AAATCACATCATGGTCTCTGCGAT 1497
1501 AATCTTTGTTAATTAAGTGGGCAAGTAGACTTCTCTGTTTGTCTTTTCTTTTCTTTT 1560
1498 CAGC-TTCTGAAGTCTGGGACATAGGAAGTTCATCTCTGTCGATCTTGAGGCTC 1556
1561 TCTTTTGTAGCTTAATGTAGATATCTTTATCATCTCTGAAATGTAATATATATTAAG 1620
1557 AGATGCTGATACC-AGTAACATTCAAAATTAACCTTCACGCAATGGAATACATTTAAAC 1612
1621 TGCTCAATTAATAGAAATGATGATGTAATTTGGATGTAATATTC--AGTTTATATAATTA 1678
1613 TGTTCAATAGCGAGTGATATTTGGACATAAATATCTCTCTTCCAAATTTTATGTAATTA 1672
1679 TATCTAATTTGTACCCCTTGTGAAATTTGTCATTTATACAAATAAGGCAATTTCTTTATCTC 1738
1673 TACCTAATCTTATACCCCTTGTAAACTTGTCACTATGCAATTAAGCAGTTGATTTTCTG 1732
1739 TAAATATGAAAAA----- 1762
1733 TGGAAAAA----- 1756

RESULT 8

ID AAH20423 standard; cDNA; 939 BP.
AC AAH20423;
DT 06-AUG-2001 (first entry)
XX Human sphingomyelinase-like phosphodiesterase 21 cDNA.
DE Sphingomyelinase-like phosphodiesterase 21; malignant tumour; haemopathy;
KW HIV infection; immunological disease; inflammation; human; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 68..631
FT /tag= a
FT /product= "Sphingomyelinase-like phosphodiesterase 21"
XX
PN WO200131030-A1.
XX
PD 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-CN00386.
XX
XX 28-OCT-1999; 99CN-0119899.
PR (SHAN-) SHANGHAI BIO ROAD GENE DEV LTD.
PA Mao Y, Xie Y;
XX
XX WPI; 2001-300511/31.
DR P-PSDB; AAB97255.
XX
XX New human acid sphingomyelinase-like phosphodiesterase 21 for diagnosis
PT and treatment of malignant tumor, hemopathy, human immunodeficiency
PT virus infection, immunological diseases and various inflammation
XX
PS Claim 6; Page 24; 29pp; Chinese.
XX

CC This invention relates to a human sphingomyelinase-like phosphodiesterase
CC 21 protein and nucleotide sequence. The invention includes a vector
CC containing the nucleotide sequence, a host cell transformed with the
CC vector, and an antibody that specifically binds to the protein. The
CC polypeptide and encoded polynucleotide are applicable in diagnosis and
CC treatment of malignant tumours, haemopathy, HIV infection, immunological
CC diseases and inflammation. The present sequence represents cDNA encoding
CC human sphingomyelinase-like phosphodiesterase 21.
XX
SQ Sequence 939 BP; 300 A; 140 C; 170 G; 329 T; 0 other;
Query Match 52.6%; Score 927; DB 22; Length 939;
Best Local Similarity 99.9%; Pred. No. 3.6e-197;
Matches 938; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 800 GGAGAAAGGTGATATATCATAGACATGTTCCAGTGGGGTATCTGCCATCTTTCAGAACAT 859
DB 1 GGAGAAAGGTGATATATCATAGACATGTTCCAGTGGGGTATCTGCCATCTTTCAGAACAT 60
QY 860 CACAGCAATGAGAGAAATCTATAATGAGAAATTTAGATATATTTTCAAAAAATACAGTGA 919
DB 61 CACAGCAATGAGAGAAATCTATAATGAGAAATTTAGATATATTTTCAAAAAATACAGTGA 120
QY 920 TGTCTATTGCGAGCAAAATTTTATGAGACACACTCACAGAGACAGCATTTATGGTTCTTTT 979
DB 121 TGTCTATTGCGAGCAAAATTTTATGAGACACACTCACAGAGACAGCATTTATGGTTCTTTT 180
QY 980 TAAAAAGGAGTCCAGTAAATCTTTTGTGTGCTCTCTGTGTACACAGTGAAGAG 1039
DB 181 TAAAAAGGAGTCCAGTAAATCTTTTGTGTGCTCTCTGTGTACACAGTGAAGAG 240
QY 1040 TGTTTTAAAAAAGAGCAACCAATCTTGTGTATACAGCTGTTTTCAGTATGATCTCTCTG 1099
DB 241 TGTTTTAAAAAAGAGCAACCAATCTTGTGTATACAGCTGTTTTCAGTATGATCTCTCTG 300
QY 1100 TTATAAATTTATTTGGATATGTTGTCAGTATTTTGAATCTGACAGAGGCGAAATCTAAAGG 1159
DB 301 TTATAAATTTATTTGGATATGTTGTCAGTATTTTGAATCTGACAGAGGCGAAATCTAAAGG 360
QY 1160 AGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGACCTACGACATTTGAAGATTGCA 1219
DB 361 AGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGACCTACGACATTTGAAGATTGCA 420
QY 1220 GCGGAAAGTTTATATGATTTAGTAAACAAATTTACAAATCTCTAGACAGTAAAGCAGTTTAT 1279
DB 421 GCGGAAAGTTTATATGATTTAGCTTAAACAAATTTACAAATCTCTAGACAGTAAAGCAGTTTAT 480
QY 1280 AAAATATACAAATTTACTTTTGTGAGTTTATGACAGAGTGTAAACATGTGATAAGACATG 1339
DB 481 AAAATATACAAATTTACTTTTGTGAGTTTATGACAGAGTGTAAACATGTGATAAGACATG 540
QY 1340 TAAGGCTTTTACAGATTTGTCGAATTTATGAAATCTTGAATATTTTCTATGCGAGATTGCT 1399
DB 541 TAAGGCTTTTACAGATTTGTCGAATTTATGAAATCTTGAATATTTTCTATGCGAGATTGCT 600
QY 1400 CAAACAGCTTTATATAAGACCAATTTACTAGTATTTTTCAGATTTTGTCTAAATAGAAAATG 1459
DB 601 CAAACAGCTTTATATAAGACCAATTTACTAGTATTTTTCAGATTTTGTCTAAATAGAAAATG 660
QY 1460 CTGATTTCTGATTTCTGAGATCAATTTTGGGAAATTTTACATAAATCTTTCTTAATTA 1519
DB 661 CTGATTTCTGATTTCTGAGATCAATTTTGGGAAATTTTACATAAATCTTTCTTAATTA 720
QY 1520 GTGGGCAAGTAGACTTCTGCTCTTCTGCTTTC-TTTTTTTTTTTCTTTTGTGCTCTTAAT 1578
DB 721 GTGGGCAAGTAGACTTCTGCTCTTCTGCTTTC-TTTTTTTTTTTCTTTTGTGCTCTTAAT 780
QY 1579 GTAGATATCTTTTATCAATCTGAAATTTGATATATATATATTTTAAAGTGTCTTAATAGAA 1638
DB 781 GTAGATATCTTTTATCAATCTGAAATTTGATATATATATTTTAAAGTGTCTTAATAGAA 840
QY 1639 TCGATGTAAATTTGGATGTAAATTTTTCAGTTTATATAAATTTATATCTTAATTTTGACCTTGT 1698

Db 781 ATTATATCTAAATTTGACCTTTGTTGAAATTTGTCATTATTAACAATAAGCGAATTCCTTTA 840
Qy 1735 TCTCTAAATATGAAAAA 1757
Db 841 TCTCTAAAAA 863
RESULT 10
ABL61985
ID ABL61985 standard; DNA; 863 BP.
XX
XX ABL61985;
XX 15-MAY-2002 (first entry)
XX Colon adenocarcinoma related gene sequence SEQ ID NO:322.
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; db.
XX Homo sapiens.
XX WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US10838.
XX 05-JUN-2000; 2000US-209473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233617P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234924P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 25-SEP-2000; 2000US-235280P.
XX 26-SEP-2000; 2000US-235637P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.
XX 28-SEP-2000; 2000US-236028P.
XX 28-SEP-2000; 2000US-236032P.
XX 28-SEP-2000; 2000US-236033P.
XX 28-SEP-2000; 2000US-236034P.
XX 28-SEP-2000; 2000US-236109P.
XX 28-SEP-2000; 2000US-236111P.
XX 29-SEP-2000; 2000US-236842P.
XX 29-SEP-2000; 2000US-236891P.
XX 02-OCT-2000; 2000US-237172P.
XX 02-OCT-2000; 2000US-237173P.
XX 02-OCT-2000; 2000US-237278P.
XX 02-OCT-2000; 2000US-237294P.
XX 02-OCT-2000; 2000US-237295P.
XX 02-OCT-2000; 2000US-237316P.
XX 03-OCT-2000; 2000US-237425P.
XX 03-OCT-2000; 2000US-237598P.
XX 03-OCT-2000; 2000US-237604P.
XX 03-OCT-2000; 2000US-237606P.
XX 03-OCT-2000; 2000US-237608P.
XX 01-NOV-2000; 2000US-244867P.
XX 01-NOV-2000; 2000US-245084P.

PA (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 322; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (1) of a signature gene set, where (1)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX Sequence 863 BP; 285 A; 125 C; 148 G; 305 T; 0 other;
SQ
Query Match 48.6%; Score 856.6; DB 24; Length 863;
Best Local Similarity 99.5%; Pred. No. 1.8e-181;
Matches 859; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 895 TAGATATTTTCAAAAATACAGTGATGTCATTGCGAGGCAATTTTATGGACACACTCACA 954
Db 1 TAGATATTTTCAAAAATACAGTGATGTCATTGCGAGGCAATTTTATGGACACACTCACA 60
Qy 955 GAGACAGCATTTATGGTTCTTTTCAGATAAAAAAGGAAGTCCAGTAATTTCTTTTGTGTGG 1014
Db 61 GAGACAGCATTTATGGTTCTTTTCAGATAAAAAAGGAAGTCCAGTAATTTCTTTTGTGTGG 120
Qy 1015 CTCCTGCTGTACACAGTGAGAGTGTTTAAAGAAAGAGGAGTCCATCTGGAAGTATATCTCTGATCA 1074
Db 121 CTCCTGCTGTACACAGTGAGAGTGTTTAAAGAAAGAGGAGTCCATCTGGAAGTATATCTCTGATCA 180
Qy 1075 GACTGTTTCAGTATGATCCTCGTATTATAAATTTTGGATATTTGTCAGTATTACTTGA 1134
Db 181 GACTGTTTCAGTATGATCCTCGTATTATAAATTTTGGATATTTGTCAGTATTACTTGA 240
Qy 1135 ATCTGACAGAGGCGAATCTAAAGGAGAGTCCATCTGGAAGTGGAGTATATCTCTGACCC 1194
Db 241 ATCTGACAGAGGCGAATCTAAAGGAGAGTCCATCTGGAAGTGGAGTATATCTCTGACCC 300
Qy 1195 AGACCTAGACATTGAAGATTTGACGCGGAAAGTTTATATGATAGTAAACAATTTA 1254
Db 301 AGACCTAGACATTGAAGATTTGACGCGGAAAGTTTATATGATAGTAAACAATTTA 360
Qy 1255 CAATCTAGACAGTAAGCAGTTTATAAATACTACAATTTCTTTTGGAGTTATGACA 1314
Db 361 CAATCTAGACAGTAAGCAGTTTATAAATACTACAATTTCTTTTGGAGTTATGACA 420
Qy 1315 GCAGTGTAACATGTGATAAGACATGTAAAGGCCCTTTTCAGATTTTGTGCAATTTATGAATCTTTG 1374
Db 421 GCAGTGTAACATGTGATAAGACATGTAAAGGCCCTTTTCAGATTTTGTGCAATTTATGAATCTTTG 480
Qy 1375 ATAATATTTCTATGAGATTGCGCTCAACACGCTTTTATATAAGCACAATTACTAGTATT 1434
Db 481 ATAATATTTCTATGAGATTGCGCTCAACACGCTTTTATATAAGCACAATTACTAGTATT 540

Db 421 CCAATCTCCAGGTTTCCCTGGCTGGGTAAATCATGACTATTGGCCAC----- 469
 Qy 549 CTGTAGTACCAGTAAAGTGTACATGAGTAACTCTGGAACCATGGCTAGAT 608
 Db 470 ----- 469
 Qy 609 GAAGAAGCTATTAGTACTTTAAGGAAGGTGGTTTTATTACAGAAAGTTACAACCTAAT 668
 Db 470 ----- 469
 Qy 669 CCAAACTTTAGGATCATCAGTCTAAACACAACTTTGTACTAGCGGCCAAATATAATGACA 728
 Db 470 ----- 469
 Qy 729 CTGAACAAGACTGACCCAGCCAAACAGTTTGAATGGCTAGAAAGTACATTGAACAACCTCT 788
 Db 470 ----- 469
 Qy 789 CAGCAGAATAAGGAAGGTGTATATCATAGCAGATGTTCCAGTGGGGTATCTGCCATCT 848
 Db 470 -----AGGTGTATATCATAGCAGATGTTCCAGTGGGGTATCTGCCATCT 513
 Qy 849 TCACAGACATCACAGCAATCAGAGAAATCTATATAGAGAAATGTAGATATTTTCAA 908
 Db 514 TCACAGAAATCACAGCAATCAGAGAAATCTATATAGAGAAATGTAGATATTTTCAA 573
 Qy 909 AAATACAGTGTATGTCAGGCAATTTATGACACACTCACAGACAGCATATTG 968
 Db 574 AAGTACAGTGTATGTCAGGCAATTTATGACACACTCACAGACAGCATATTG 633
 Qy 969 GTTCTTTTCAGATAAAGGAGTCCAGTAAATCTTTGTTGGCTCTGCTGTAC 1028
 Db 634 GTTCTTTTCAGATAAAGGAGTCCAGTAAATCTTTGTTGGCTCTGCTGTAC 693
 Qy 1029 CCAGTGAAGAGTGTTTAGAAAAACAGCAACATCTGGTATCAGACTGTTTCAGTAT 1088
 Db 694 CCAGTGAAGAGTGTTTAGAAAAACAGCAACATCTGGTATCAGACTGTTTCAGTAT 753
 Qy 1089 GATCCTCGTATTAATTTATGATATGTCAGTATTTGTCAGTATTTGTCAGTATTTGTCAGTAT 1148
 Db 754 GATCCTCGTATTAATTTATGATATGTCAGTATTTGTCAGTATTTGTCAGTATTTGTCAGTAT 813
 Qy 1149 AATCTAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCAAGACCTACACATTT 1208
 Db 814 AATCTAAGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCAAGACCTACACATTT 873
 Qy 1209 GAAGATTTGACGCGGAAAGTTTATATGATTTAGCTAAACAAATTTTACATCTTAGACAGT 1268
 Db 874 GAAGATTTGACGCGGAAAGTTTATATGATTTAGCTAAACAAATTTTACATCTTAGACAGT 933
 Qy 1269 AAGCAGTTTATAAATACATACTTCTTTGAGTATGACAGCAGTGAACATGT 1328
 Db 934 AAGCAGTTTATAAATACATACTTCTTTGAGTATGACAGCAGTGAACATGT 993
 Qy 1329 GATAAGACATGAAGGCTTTTCAGATTTGTGCAATTTATGATTTGATATTTTCTAT 1388
 Db 994 GATAAGACATGAAGGCTTTTCAGATTTGTGCAATTTATGATTTGATATTTTCTAT 1053
 Qy 1389 GCAGATTCCTCAACAGCTTTTATATAAGCACAATTTACTAG 1430
 Db 1054 GCAGATTCCTCAACAGCTTTTATATAAGCACAATTTACTAG 1095

RESULT 13
 ID AAA02374
 XX AA02374 standard; cDNA; 728 BP.
 AC AAA02374;
 XX
 DT 19-MAY-2000 (first entry)
 XX
 DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2365.

XX Human; colon cancer; tumour; diagnosis; gene expression product;
 KW probe; detection; cancerous state; metastasis; identification;
 KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
 KW oestrogen receptor-negative breast cancer; lung cancer; ss.
 XX Homo sapiens.
 OS
 XX
 FN WO9958675-A2.
 XX
 PD 18-NOV-1999.
 XX
 PF 13-MAY-1999; 99WO-US10602.
 XX
 PR 14-MAY-1998; 98US-0085426.
 PR 15-MAY-1998; 98US-0085537.
 PR 15-MAY-1998; 98US-0085696.
 PR 21-OCT-1998; 98US-0105234.
 PR 27-OCT-1998; 98US-0105877.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
 XX WPI; 2000-126369/11.
 XX
 PT Polynucleotide library used to determine cancerous states of mammalian
 cells -
 XX
 PS Claim 1; Page 940; 1097pp; English.
 XX
 CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 libraries constructed from human colon cancer cell lines. The present
 invention also describes a method of detecting differentially expressed
 genes correlated with a cancerous state of a mammalian cell, comprising
 detecting at least one differentially expressed gene product in a test
 sample derived from a cell suspected of being cancerous, where detection
 of the differentially expressed gene product is correlated with a
 cancerous state of the cell from which the test sample was derived.
 The polynucleotides sequences can be used in a method for detecting
 differentially expressed genes correlated with a cancerous state of a
 mammalian cell. The polynucleotides can also be used as probes for
 detecting and mapping related genes. They can be used in diagnosis and
 prognosis of diseases and disorders (e.g. identification of
 pre-metastatic or metastatic cancerous states, stages of cancer, or
 responsiveness of cancer to therapy). This is particularly for breast
 cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 negative breast cancer, lung cancer, and colon cancer.
 CC
 SQ Sequence 728 BP; 240 A; 155 C; 139 G; 189 T; 5 other;
 Query Match 39.8%; Score 701.8; DB 21; Length 728;
 Best Local Similarity 98.9%; Pred. No. 6e-147; Indels 0; Gaps 0;
 Matches 706; Conservative 0; Mismatches 8;
 Qy 383 GATATGGACAGGGATAGCCACCTCATGTTCTGTACCTGAACCTCTCAACAGACACTGT 74
 Db 15 GATTCGGACAGGAGTAGCCACCTCATGTTCTGTACCTGAACCTCTCAACAGACACTGT 442
 Qy 443 TATAAATGTGATCACTAATATGACACCAACCATCCAGAGTCTCTTTCAAAATCTCCAGT 502
 Db 75 TATAAATGTGATCACTAATATGACACCAACCATCCAGAGTCTCTTTCCAAATCTCCAGT 134
 Qy 503 TTTCCCTCGCTGGTAACTATGACTATTGGCCACAGGATCAACTGCTGTAGTCACCAG 562
 Db 135 TTTCCCTCGCTGGTAACTATGACTATTGGCCACAGGATCAACTGCTGTAGTCACCAG 194
 Qy 563 TAAAGTGTACAAATCAGTAGCAACCTCTGGAACCATGCTAGATGAAGAGCTATTAG 622

Db 195 TAAAGTGACAACTGAGTACAAACCTCTGGAAACCATGGCTAGATGAAGAAGCTATTAG 254
 Qy 623 TACTTTAAGGAAGTGGTTTTTATTCACAGAAAGTTACAACTAATCCAAACCTTAGGAT 682
 Db 255 TACTTTAAGGAAGTGGTTTTTATTCACAGAAAGTTACAACTAATCCAAACCTTAGGAT 314
 Qy 683 CATCAGTCTAAACACAACTGCTACTACGGCCCAATATAATGACACTGAAACAAGACTGA 742
 Db 315 CATCAGTCTAAACACAACTGCTACTACGGCCCAATATAATGACACTGAAACAAGACTGA 374
 Qy 743 CCCAGCCCAACGATTTGAATGCTGAGAAAGTACATTTGAACAACCTCTCAGCAGAAATAAGGA 802
 Db 375 CCCAGCCCAACGATTTGAATGCTGAGAAAGTACATTTGAACAACCTCTCAGCAGAAATAAGGA 434
 Qy 803 GNAAGTGATATCATAGACATGTTCCAGTGGGGTATCTGCCATCTTTCAGAAACATCAC 862
 Db 435 GNAAGTGATATCATAGACATGTTCCAGTGGGGTATCTGCCATCTTTCAGAAACATCAC 494
 Qy 863 AGCAATGAGAGAACTACTATAATGAGAAATTTGATAGATATTTTCAAAAATACAGTGATCT 922
 Db 495 AGCAATGAGAGAACTACTATAATGAGAAATTTGATAGATATTTTCAAAAATACAGTGATCT 554
 Qy 923 CATTCAGGACAAATTTATGACACACTCACAGAGACAGCAATATGTTCTTTTCAGATAA 982
 Db 555 CATTCAGGACAAATTTATGACACACTCACAGAGACAGCAATATGTTCTTTTCAGATAA 614
 Qy 983 AAAAGGAAGTCCAGTAAATCTTTGTTGGCTCTCTGCTGTTACACAGTGAAGAGTGT 1042
 Db 615 AAAAGGAAGTCCAGTAAATCTTTGTTGGCTCTCTGCTGTTACACAGTGAAGAGTGT 674
 Qy 1043 TTTAGAAAAACAGACCAAACTCTGGTATCAGACTGTTTTCAGTATGATCCTCG 1096
 Db 675 TTTAGAAAAACAGACCAAACTCTGGTATCAGACTGTTTTCAGTATGATCCTCG 728

RESULT 14

ID AAA02333 standard; cDNA; 733 BP.
 AC AAA02333;
 XX
 DT 19-MAY-2000 (first entry)
 XX
 DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2324.
 XX
 KW Human; colon cancer; tumour; diagnosis; gene expression product;
 KW probe; detection; cancerous state; metastasis; identification;
 KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
 KW oestrogen receptor-negative breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9598675-A2.
 XX
 PD 18-NOV-1999.
 XX
 PF 13-MAY-1999; 99WO-US10602.
 XX
 PR 14-MAY-1998; 98US-0085426.
 PR 15-MAY-1998; 98US-0085537.
 PR 15-MAY-1998; 98US-0085696.
 PR 21-OCT-1998; 98US-0105234.
 PR 27-OCT-1998; 98US-0105877.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYGE-) HYSEQ INC.
 XX
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamson G, Drmanac R, Crkvenjakov R, Drmanac M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
 XX WPI; 2000-126369/11.

XX Polynucleotide library used to determine cancerous states of mammalian cells -
 PT
 XX Claim 1; Page 919; 1097pp; English.
 XX
 CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived.
 CC The polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of
 CC pre-metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer.
 XX
 SQ Sequence 733 BP; 241 A; 158 C; 140 G; 190 T; 4 other;
 Query Match 37.3%; Score 657.4; DB 21; Length 733;
 Best Local Similarity 97.9%; Pred. No. 4.9e-137;
 Matches 698; Conservative 0; Mismatches 11; Indels 4; Gaps 3;
 Qy 388 GGACAGGGGATAGCCCACTCATGTTCTGTACCTGAACCTCTCAACAGACACTGTTATAA 447
 Db 21 GCACAGGGGATAGCCCACTCATGTTCTGTACCTGAACCTCTCAACAGACACTGTTATAA 80
 Qy 448 ATGTGATCACTAATATGACAAACCATCCAGAGTCTCTTCCAAATCTCCAGGTTTCC 507
 Db 81 ATGTGATCACTAATATGACAAACCATCCAGAGTCTCTTCCAAATCTCCAGGTTTCC 140
 Qy 508 CTGCGCTGGGTANTCATGACTATTTGGCCACAGGATCACTGCTGTAGTCACAGTAAAG 567
 Db 141 CTGCGCTGGGTANTCATGACTATTTGGCCACAGGATCACTGCTGTAGTCACAGTAAAG 200
 Qy 568 TGTCAATATGCTAGTACAAACCTCTCGAAACCATGCTAGATGAAGAAGCTATTAGTACTT 627
 Db 201 TGTCAATATGCTAGTACAAACCTCTCGAAACCATGCTAGATGAAGAAGCTATTAGTACTT 260
 Qy 628 TAAGGAAAGGTGGTTTTTATTACAGAAAGTTTCAACTAATCCAAACCTTAGATCATCA 687
 Db 261 TAAGGAAAGGTGGTTTTTATTACAGAAAGTTTCAACTAATCCAAACCTTAGATCATCA 320
 Qy 688 GTCTAAACACAACTTGTACTACGGCCCAATATAATGACACTGACAGACTGACCCAG 747
 Db 321 GTCTAAACACAACTTGTACTACGGCCCAATATAATGACACTGACAGACTGACCCAG 380
 Qy 748 CCAACAGTTTGAATGGCTAGAAAGTACATTTCAACAACTCTCAGCAGAAATAAGGAAGG 807
 Db 381 CCAACAGTTTGAATGGCTAGAAAGTACATTTCAACAACTCTCAGCAGAAATAAGGAAGG 440
 Qy 808 TGTATATCATAGACATGTTCCAGTGGGGTATCTGCCATCTTTCAGAAACATCACAGCAA 867
 Db 441 TGTATATCATAGACATGTTCCAGTGGGGTATCTGCCATCTTTCAGAAACATCACAGCAA 500
 Qy 868 TCAGAGAACTACTATAATGAGAAATTTGATAGATATTTTCAAAAATACAGTGTGCTATG 927
 Db 501 TCAGAGAACTACTATAATGAGAAATTTGATAGATATTTTCAAAAATACAGTGTGCTATG 560
 Qy 928 CAGGACAAATTTTATGGACACACTCACAGAGACAGCAATATGTTCTTTTCAGATAAAAAA- 986
 Db 561 CAGGACAAATTTTATGGACACACTCACAGAGACAGCAATATGTTCTTTTCAGATAAAAAA 620
 Qy 987 GGAAGTCCAGTAAATCTTTTGTGGTCTCTCTGCTGTGTACACC--AGTGAAGAGTGT 1044
 Db 621 GGAAGTCCAGTAAATCTTTTGTGGTCTCTCTGCTGTGTACACCAGTGAAGAGTGT 680

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:39:45 ; Search time 58.7918 Seconds
(without alignments)
9201.581 Million cell updates/sec

Title: US-09-981-353-47
Perfect score: 1764
Sequence: 1 ctcaggcctgacggtccgag.....tgaaaaaaaaaaaaagg 1764

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1764	100.0	1764	3	US-09-013-881-13
2	54.8	3.1	5852	1	US-07-867-106-2
3	54	3.1	658	4	US-08-998-416-595
4	51.4	2.9	2058	2	US-08-749-391-1
5	51.4	2.9	2058	3	US-09-390-200-1
C 6	48.8	2.8	3701	4	US-08-845-258-10
C 7	48.8	2.8	3701	4	US-08-990-571-10
C 8	48.8	2.8	3701	4	US-08-723-142A-10
C 9	48.8	2.8	3701	4	US-09-528-784A-10
C 10	48.8	2.8	19124	2	US-08-487-826B-13
C 11	48.6	2.8	8920	4	US-08-446-855A-1
C 12	48.6	2.8	8920	4	US-09-150-741-1
C 13	48.4	2.7	635	4	US-08-998-416-1137
C 14	48.4	2.7	2447	2	US-09-014-969-14
C 15	48.4	2.7	2852	3	US-09-027-137-2
C 16	48.4	2.7	2852	4	US-09-344-441-2
C 17	45.4	2.6	4673	1	US-07-638-431-1
C 18	45.4	2.6	4673	5	PCT-US92-00018-1
C 19	45.4	2.6	19124	2	US-08-487-826B-13
C 20	44.6	2.5	6243	2	US-09-056-075-1
C 21	44.6	2.5	20674	4	US-09-641-638-651
C 22	43.8	2.5	240	1	US-08-628-417-6
C 23	43.8	2.5	837	4	US-08-998-416-288
C 24	43.8	2.5	3865	1	US-08-832-883-48
C 25	43.8	2.5	3865	2	US-08-832-877-48
C 26	43.6	2.5	3581	2	US-08-738-349-1
C 27	43.4	2.5	20674	4	US-09-641-638-651

28	43.2	2.4	6152	4	US-08-973-462-1	Sequence 1, Appli
29	43	2.4	602	1	US-08-764-100-8	Sequence 8, Appli
C 30	43	2.4	642	1	US-08-764-100-13	Sequence 13, Appli
31	43	2.4	643	1	US-08-764-100-7	Sequence 7, Appli
C 32	43	2.4	1441	4	US-08-821-994-63	Sequence 63, Appli
33	43	2.4	2993	1	US-08-764-100-2	Sequence 2, Appli
C 34	43	2.4	2993	1	US-08-764-100-10	Sequence 10, Appli
C 35	43	2.4	3000	1	US-08-764-100-9	Sequence 9, Appli
36	43	2.4	3001	1	US-08-764-100-1	Sequence 1, Appli
37	43	2.4	4818	3	US-08-817-926-27	Sequence 27, Appli
C 38	43	2.4	6124	4	US-08-213-419B-3	Sequence 3, Appli
C 39	43	2.4	11485	4	US-09-410-464-9	Sequence 9, Appli
40	42.8	2.4	3701	4	US-08-845-258-10	Sequence 10, Appli
41	42.8	2.4	3701	4	US-08-990-571-10	Sequence 10, Appli
42	42.8	2.4	3701	4	US-08-723-142A-10	Sequence 10, Appli
43	42.8	2.4	3701	4	US-09-528-784A-10	Sequence 10, Appli
44	42.8	2.4	5923	4	US-09-064-922-3	Sequence 3, Appli
C 45	42.6	2.4	935	2	US-08-892-770-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-013-881-13
; Sequence 13, Application US/09013881
; Patent No. 6132964
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,881
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0470 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGNOT10
; CLONE: 1376382
; US-09-013-881-13

Query Match 100.0%; Score 1764; DB 3; Length 1764;									
Best Local Similarity 100.0%; Pred No. 0;									
Matches 1764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CTCAGGCTGACGGTCCGAGTGGAGTGGGGACAGCCGAGAACCTCCAGGTGAGCCCGC	60						
DB	1	CTCAGGCTGACGGTCCGAGTGGAGTGGGGACAGCCGAGAACCTCCAGGTGAGCCCGC	60						
QY	61	GGCCCTCCATGGGCTGGTGGGGACATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	120						
DB	61	GGCCCTCCATGGGCTGGTGGGGACATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	120						
QY	121	GCTCCGGCTCCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	180						
DB	121	GCTCCGGCTCCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	180						
QY	181	GACAGTTTGGCATGTGACTGACTTACACTTACAGCTTACAGCTTACAGCTTACAGCTTACAG	240						
DB	181	GACAGTTTGGCATGTGACTGACTTACACTTACAGCTTACAGCTTACAGCTTACAGCTTACAG	240						
QY	241	ACACAAAGTGTGCTTTCATCTAAAGTGCAGTGCCTTCCAACTTCCAACTTCCAACTTCCAA	300						
DB	241	ACACAAAGTGTGCTTTCATCTAAAGTGCAGTGCCTTCCAACTTCCAACTTCCAACTTCCAA	300						
QY	301	ATGTTCTGTGATCTCCATATCAACTTATTTTGTGAGCTTATTTTGTGAGCTTATTTTGTGAG	360						
DB	301	ATGTTCTGTGATCTCCATATCAACTTATTTTGTGAGCTTATTTTGTGAGCTTATTTTGTGAG	360						
QY	361	CTGGACAAAGCATCTTTTCATGATATGGACGGGATAGCCACCTCATGTTCTCTGTAC	420						
DB	361	CTGGACAAAGCATCTTTTCATGATATGGACGGGATAGCCACCTCATGTTCTCTGTAC	420						
QY	421	CTGAACCTCAACAGACCTGTTTAAATGTGATCACTAATATGACAAACCATCCAG	480						
DB	421	CTGAACCTCAACAGACCTGTTTAAATGTGATCACTAATATGACAAACCATCCAG	480						
QY	481	GTCCTTTTCCAAATCTCCAGTGTTCCTCGCTGGTAAATCATGACTTATGAGCCACAGG	540						
DB	481	GTCCTTTTCCAAATCTCCAGTGTTCCTCGCTGGTAAATCATGACTTATGAGCCACAGG	540						
QY	541	ATCAACTGCTGTAGTCAACAGTAAAGTGTACAAATGACAGTAAACCTCTGGAACCAT	600						
DB	541	ATCAACTGCTGTAGTCAACAGTAAAGTGTACAAATGACAGTAAACCTCTGGAACCAT	600						
QY	601	GGCTAGATGAAGAGCTATTAGTACTTTAAGGAAAGGTGGTTTTTATTCACAGAAAGTTA	660						
DB	601	GGCTAGATGAAGAGCTATTAGTACTTTAAGGAAAGGTGGTTTTTATTCACAGAAAGTTA	660						
QY	661	CAACTAATCCAACTTAGATCATCAGTCTAAACAACTTGTACTACGGCCCAATA	720						
DB	661	CAACTAATCCAACTTAGATCATCAGTCTAAACAACTTGTACTACGGCCCAATA	720						
QY	721	TAATGACACTGAACAGTACCCAGCCCAACAGTTTGAATGGCTAGAAAGTACATTGA	780						
DB	721	TAATGACACTGAACAGTACCCAGCCCAACAGTTTGAATGGCTAGAAAGTACATTGA	780						
QY	781	ACAACTCTCAGCAGATAGGAGAGGTGTATATCATAGCACATGTTCCAGTGGGTATC	840						
DB	781	ACAACTCTCAGCAGATAGGAGAGGTGTATATCATAGCACATGTTCCAGTGGGTATC	840						
QY	841	TGCCATCTTTCACAGAACTCAGCAATGAGAGTACTTATATGAGAAATTTGATAGATA	900						
DB	841	TGCCATCTTTCACAGAACTCAGCAATGAGAGTACTTATATGAGAAATTTGATAGATA	900						
QY	901	TTTTTCAAAATACAGTGTGCTTTCAGTGGGCAATTTTATGGACACTCAGAGACA	960						
DB	901	TTTTTCAAAATACAGTGTGCTTTCAGTGGGCAATTTTATGGACACTCAGAGACA	960						
QY	961	GCATTATGTTCTTTCAGATATAAAGGAGTCCAGTAAATTTCTTTTGTGGTCTCTG	1020						
DB	961	GCATTATGTTCTTTCAGATATAAAGGAGTCCAGTAAATTTCTTTTGTGGTCTCTG	1020						

QY	1021	CTGTTACACCACTGAAGAGTGTGTTTAGAAAAACAGACCAACAACTCTGCTATCAGACTGT	1080
DB	1021	CTGTTACACCACTGAAGAGTGTGTTTAGAAAAACAGACCAACAACTCTGCTATCAGACTGT	1080
QY	1081	TTTCAGTATGATCCTCGTGATTAATAATTATTTGGATATGTTGGCAGTATTACTTTGAATCTGA	1140
DB	1081	TTTCAGTATGATCCTCGTGATTAATAATTATTTGGATATGTTGGCAGTATTACTTTGAATCTGA	1140
QY	1141	CAGAGCGCAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGACCT	1200
DB	1141	CAGAGCGCAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACCCAGACCT	1200
QY	1201	ACGACATTTGAAGATTTCAGCGCGGAAAGTTTATATGGATTAGCTAAACAAATTTACAATCC	1260
DB	1201	ACGACATTTGAAGATTTCAGCGCGGAAAGTTTATATGGATTAGCTAAACAAATTTACAATCC	1260
QY	1261	TAGACAGTAAGCAGTTTATAAATACTACAATTACTTCTTTGTGAGTTATGACAGCAGTG	1320
DB	1261	TAGACAGTAAGCAGTTTATAAATACTACAATTACTTCTTTGTGAGTTATGACAGCAGTG	1320
QY	1321	TAAACATGTGATAAGACATGTAAGGCTTTTCAAGTTTGTGCAATTTATGAATCTTGATAATA	1380
DB	1321	TAAACATGTGATAAGACATGTAAGGCTTTTCAAGTTTGTGCAATTTATGAATCTTGATAATA	1380
QY	1381	TTTCTCTATGCAATTTGCTTCAAAACAGCTTTTATATAAGCACAAATTTACTAGTATTTACACAG	1440
DB	1381	TTTCTCTATGCAATTTGCTTCAAAACAGCTTTTATATAAGCACAAATTTACTAGTATTTACACAG	1440
QY	1441	TTTTTGTCTAAATAGAAAATGCTGATTTCTGATTTCTGATGATCAATTTGTGGGAATTTTACATA	1500
DB	1441	TTTTTGTCTAAATAGAAAATGCTGATTTCTGATTTCTGATGATCAATTTGTGGGAATTTTACATA	1500
QY	1501	AATCTTTGTTAAATGCTGAGTGGCAAGTGTGCTTTCTGCTTTCTTTTCTTTTCTTTTCTTTT	1560
DB	1501	AATCTTTGTTAAATGCTGAGTGGCAAGTGTGCTTTCTGCTTTCTTTTCTTTTCTTTTCTTTT	1560
QY	1561	TCCTTTTGTGATGCTTTAAATGATAGATATCTTTTATCTTGAATTTGATATATATTTAAAG	1620
DB	1561	TCCTTTTGTGATGCTTTAAATGATAGATATCTTTTATCTTGAATTTGATATATATTTAAAG	1620
QY	1621	TGCTCATTTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1680
DB	1621	TGCTCATTTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1680
QY	1681	TCTAATTTGTACCTTGTGTAATTTGTCAATTTATATATATATATATATATATATATATATAT	1740
DB	1681	TCTAATTTGTACCTTGTGTAATTTGTCAATTTATATATATATATATATATATATATATATAT	1740
QY	1741	AATATCAAAAAAAGG 1764	
DB	1741	AATATCAAAAAAAGG 1764	

RESULT 2

US-07-867-106-2

Sequence 2, Application US/07867106

Patent No. 5389526

GENERAL INFORMATION:

APPLICANT: Slade, Martin B

APPLICANT: Chang, Andy C M

APPLICANT: Williams, Keith L

TITLE OF INVENTION: Improved plasmid Vectors for Cellular

TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSES: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris

STREET: One Liberty Place 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT: US/07/867,106
FILING DATE: 19920625
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3439
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match 3.1%; Score 54.8; DB 1; Length 5852;
Best Local Similarity 45.7%; Pred. No. 0.00037;
Matches 191; Conservative 0; Mismatches 227; Indels 0; Gaps 0;
Qy 1345 CCTTTCAGATTGTGCAATATGTAATCTTGATAATATTTCTATGCGAGATTGCTCAAC 1404
Db 1620 CCCCTCTTTTTTTTTTTTTTTTGTGCATGACACTTTTTTTTTTTTGTGCATGACACTTTT 1679
Qy 1405 AGCTTTATATAGACCAATCTAGTATTTTCACAGTTTTTGCTAATAGAAATGCTGAT 1464
Db 1680 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1739
Qy 1465 TCTGATCTGAGATCAATTTGTGGGAATTTTACATAAATCTTTGTTAATTACTGAGTGG 1524
Db 1740 TTTTGTAGATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1799
Qy 1525 CAAGTAGACTTCCTGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1584
Db 1800 TGAATAATATTTAAATTTTAAATAAATCTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTAA 1859
Qy 1585 ATCTTTATCTGATGTTGATTTATATATTTAAAGTCTCATTAATAGATGATGATG 1644
Db 1860 ATTAAAAATTTTTTTTTTTTATAGATCTCATAATTAATAAATCAATTTAAAAATTAAG 1919
Qy 1645 TAAATCGATGTAATATTCAGTTTATATATTAATTAATCTAATTTGTACCCTTGTGTAAT 1704
Db 1920 TTATTTTAAATATGCAAAACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1979
Qy 1705 TGTCAATTTACAAATAAGGCAATCTTTATCTCTAAATATGAAAAAATAAATAAATAAATAA 1762
Db 1980 TTTCTTTTTTTTTTTTTTTTTTTTTTTTCTTTTGAATAAATAAATAAATAAATAAATAAATAA 2037

RESULT 3
US-08-998-416-595
Sequence 595, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippen, Peter
APPLICANT: Pohlmann, Rainer

APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 595:
SEQUENCE CHARACTERISTICS:
LENGTH: 658 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1408RP
US-08-998-416-595

Query Match 3.1%; Score 54; DB 4; Length 658;
Best Local Similarity 48.1%; Pred. No. 0.0002;
Matches 133; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
Qy 1428 TAGTATTTCCACAGTTTTTGTCTAATAGAAAATGCTGATTTCTGAGATCAATTTGTG 1487
Db 267 TATTAATTTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 336
Qy 1488 GGAATTTTACATAAATCTTTGTTTAATTAATGAGTGGGCAAGTAGACTTCTGCTTTGCT 1547
Db 327 ATTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 386
Qy 1548 TTTCTTTTTTTTTTTCTTTTGTGCTTAATGAGATATCTTTATCATTTCTGAAATTTGAT 1607
Db 387 TTATTAACCTTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 446
Qy 1608 TATATATTTAAAGTGTCTAATTAAGATGATGATGATAATTTGGAATTTGGAATTTTCAGT 1667
Db 447 TATATTTATCATGATAGTATAGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 506
Qy 1668 TTATATAATATATCTAAATTTGTACCCTTTGTGAATTTGTCAATTTATACATAAAGCAA 1727
Db 507 TGTGAACATAATATATATGCTATTTCTATTTTATTTATTTATTTATTTATTTATTTATTT 566
Qy 1728 TTCTTTATCTCTAAATAT 1745
Db 567 TTTATTTTATTTATTTT 584

US-08-487-826B-13

```
Query Match          2.8%; Score 48.8; DB 2; Length 19124;
Best Local Similarity 47.4%; Pred. No. 0.022; Indels 6; Gaps 1;
Matches 183; Conservative 0; Mismatches 197;

QY 1336 CATGTAAGCCCTTTCAGATTTGTGCAATTAATGAAATCTTGATAATATTCCTATGAGATT 1395
DB 15975 CATGTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15916

QY 1396 GCCTCAACAGCTTTATATAAGACAAATTAATGAAATCTTGATAATATTCCTATGAGATT 1455
DB 15915 TCATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15856

QY 1456 AATGCTGATCTGATCTTCAGATCTTCTGCTTTGCTTTCTTTTCTTTTCTTTTCTTTTCTTT 1515
DB 15855 TATTTAATTTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15796

QY 1516 CTGAGTGGCAAGTAGACTTCTGCTTTGCTTTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1575
DB 15795 ATGTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15736

QY 1576 AATGTAGATATCTTTA-----TCATCTGAAATGTAATATTTTAAAGTGTCTATTA 1629
DB 15735 TATTTTAAATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 15676

QY 1630 ATAGAATGATGATGTAATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1689
DB 15675 ATGTTTTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 15616

QY 1690 TACCTGTTGCAATTTGCAATTTA 1715
DB 15615 AATTTTTTTTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 15590
```

RESULT 11

```
US-08-446-855A-1/c
; Sequence 1, Application US/08446855A
; Patent No. 5849573
; GENERAL INFORMATION:
; APPLICANT: Stewart, Thomas S
; APPLICANT: Flores, Maria V
; APPLICANT: O'Sullivan, William J
; TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
; TITLE OF INVENTION: phosphate synthetase II
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nixon & Vanderhye PC
; STREET: 1100 No. 5849573th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,855A
; FILING DATE: 06-Jul-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Micchard, Leonard C
; REGISTRATION NUMBER: 29.009
; REFERENCE/DOCKET NUMBER: 47-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8920 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic
US-08-446-855A-1
```

```
Query Match          2.8%; Score 48.6; DB 2; Length 8920;
Best Local Similarity 43.7%; Pred. No. 0.017;
Matches 213; Conservative 0; Mismatches 274; Indels 0; Gaps 0;
```

```
QY 1278 ATAAATACCTTACAAATTAATTTGAGTATTAAGACAGAGTGTACATGTGATAGACA 1337
DB 1015 AAAATTTGAATAACTGTTTATACCTCTTTTATAAAAAAATAATTAATTAATAAAA 956

QY 1338 TGTAAGCCCTTTTCAGATTTGTCGAATTAATGAAATCTTTGATTAATTTCTATGAGATTGC 1397
DB 955 TAGGGATTCCTTACAAATGATAAAATGAAATATACAGAAATATATATAGAAATATAGT 896

QY 1398 CTCAAAACAGCTTTATATAAAGCACAAATTAATGAAATCTTTGCTAAAGAAA 1457
DB 895 TTTATATAGAACCAATTAAGATATATATACCTTTTAAATAACAACCTTTGTGATGTTAAAGA 836

QY 1458 TGTGATCTGATCTGAGATCAATTTGTGGAAATTTTACATAAAATCTTTGTTAAATTAAT 1517
DB 835 ATAAAACTGTTTAAAGACCTTATGATTCAGAGAAATATCCCAATTAATATATATATATAT 776

QY 1518 GAGTGGCAAGTAGACTTCTGCTTTGCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1577
DB 775 ATTTATATATATATCTATATATTTTCCCACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTAT 716

QY 1578 TGTAGATATCTTTATCATCTGAAATGTAATATATATATATTAAGTCTCATTAATAGAAATG 1637
DB 715 ATTTATGTTTAAATATTTTATAAATTTACATATACAGTTCATTTTTCATATGTAATTTT 656

QY 1638 ATGATGTAATTTGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1697
DB 655 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 596

QY 1698 TTGAAATTTGCTATTTATACATAAAGCAATCTTTATCTCTCTAAATATGAAAAAATAAAA 1757
DB 595 TATAAATCAAGAAAAAATAAATAAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 536

QY 1758 AAAAAGG 1764
DB 535 TTATAGG 529
```

RESULT 12

```
US-09-150-741-1/c
; Sequence 1, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; Patent No. 6183996
; TITLE OF INVENTION: Synthetase II
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PL6380
; EARLIER FILING DATE: 1992-12-16
; EARLIER APPLICATION NUMBER: AU93/00617
; EARLIER FILING DATE: 1993-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8920
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
US-09-150-741-1
```

Query Match

2.8%; Score 48.6; DB 4; Length 8920;

	Best Local Similarity	43.7%;	Pred. No.	0.017;	
Matches	213; Conservative	0;	Mismatches	274;	Indels
	Gaps	-0;			
QY	1278	ATAAAACTACTACAATTACTTCTTTGTGAGTATGTACACAGCAGTGAACATGTGATAAGACA	1337		
Ddb	1015	AAAATTGAAATAACTGTTTTATACCCTTTTTATAAABAAAABAATAATAATTATAATAAA	956		
QY	1338	TGTAAGGCCTTCAGATTTGTGGCAATATGGAATCTTGATAAATTTCTCTATGCGAGATGC	1397		
Ddb	955	TAGGGATGCTTACAANAATGATAAANAATGAAATATACAGAATATATTAGAACTACTATAGT	896		
QY	1398	CTCAAAACAGCTTTATATAAAGCACAAATTACTAGTATTTTCACAGTTTTTGTCTAATAGAAAA	1457		
Ddb	895	TTTATATAGAACCAATTAAGATATATATACCTTTTATAACAACTTTGTGATGTTAAAGA	836		
QY	1458	TGCTGATTCGTATTCAGATCAATTTGTGGGAATTTTACATAAACTTTGTTAAATTACT	1517		
Ddb	835	ATAAAACTGTTTAAAGACCTATGATTCAGAGAAATATCCCAATAATTATATATATATAT	776		
QY	1518	GAGTGGGCAAGTAGACTTCCTGTCTTGTGCTTCTTTTTTTTTTTTCTTTTTTGATGCCCTAA	1577		
Ddb	775	ATTTATATATATATCATATATATATTTTTTCCCATTCTTTTTTTTTTTTATACATTTAT	716		
QY	1578	TGTAGATATCTTTATCATCTCGAATTGATATATATATATTTAAAGTGCTCAITTAATAGAAATG	1637		
Ddb	715	ATTTATGTTTAAATATTTTATAATTTTACATATACAGTTCAITTTTCATATGTAAATTT	656		
QY	1638	ATGGATGTAAATTGGAGTGAATAATTACAGTTTTATATAATTAATATCTAAATTTGTACCCCTG	1697		
Ddb	655	TTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTAGTAGAAATTACTATTT	596		
QY	1698	TTGAAATGTCTTTATACAATAAAGCGAATCTCTTATCTCTAAATGATGAAAAAATAAAA	1757		
Ddb	595	TATAAACTAAGAAAAAANAATTAATTAATGAATAAATTAATTAATTAATTAATTAATAA	536		
QY	1758	AAAAAGG	1764		
Ddb	535	TTATAGG	529		

```

RESULT 13
US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgan
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebeschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF THE HUMAN
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corpore
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998, 416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97

```

FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGCJ1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1137:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1692RP

US-08-998-416--1137

Query Match 2.7%; Score 48.4; DB 4; Length 636;
Best Local Similarity 49.2%; Pred. No. 0.0053;
Matches 127; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 1434 TTCACAGTTTTTCTAATGAGAAATGCTGATTCGTGATTCGAGATCAATTTGGGAATT 1493
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 TTTAAATTTTAAATATTGTAATAATTATTATTTTATTATAATATCTATTTTATAAATA 419
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1494 TTACATAAATCTTTTGTTAATTACTGAGTGGCGAAGTAGACTTCCTGCTTTGCTTTCTTT 1553
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 TTAATGTTGATTTATATATTATTAATCTTTTATAAGAATAATTATTATTAATAATTAATTTAAC 479
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1554 TTTTTCCTCTTTTGATGCCTTAATGTAGATATCTTTATCATTTCTGAAATGTATTATATA 1613
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 TTTAAATCTTATTATTAATTTTATATTATTATAAATAATATTTATTTATTTATTT 539
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1614 TTTAAAGTGCTCAATTAATGAAATGATGGATGTAATTTGGAATGTAATTTACGTTTATAT 1673
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 ATTTATTTAAATTAATAATTAATTTATTTAAATAATTTTATCATTTATTAAATTAATAATA 599
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1674 AATTATATCTAATTTCTGA 1691
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 AATATTATAAGAATGTA 617
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-014-969-14/c
Sequence 14, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallie, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

TELEPHONE: 650-955-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 2852 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSNOT16
CLONE: 2229466
US-09-027-137-2

```

Query Match	2.7%;	Score 48.4;	DB 3;	Length 2852;
Best Local Similarity	50.9%;	Pred. No. 0.011;		
Matches 115;	Conservative 0;	Mismatches 111;	Indels 0;	Gaps 0;
QY	1534	TTCTGTGCTTTGCTTTCTTTTTTTTTTTTCTTTTTCATGCGCTAAATCTAGATATCTTTATC	1593	
Db	2584	TT	2525	
QY	1594	ATTCGAATTGTATTATATATTTAAAGTGCTCATTAATAGAAATGATGGATGTAATAATTGGA	1653	
Db	2524	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGCGAGTGTAAAAAGAAATCTTTATTTTACAAAAT	2465	
QY	1654	TGTAATAATTTCAGTTTATATAATATATATCTAAATTTGTACCCCTTGTGGAAATTCGTCAATTA	1713	
Db	2464	TAAACAATAAATAGTATTTTACAGCATCTTAAATACCAAGCTTATATAACTTGACATTTT	2405	
QY	1714	TTCAATAAAGCGAAATCTTTTATCTCTPAAATATAGAAAAAATAAAAAA	1759	
Db	2404	CTAATCATTTTCAAATTTGAGTTTCCCATAGAGCTTTTAAAAATGCAAGA	2359	

Search completed: June 28, 2003, 08:19:53
Job time : 60.7918 secs

RESULT 15

1245	Db	ACGACATTGAAGATTTCGACCGCGAAAGTTTTATATGGATTAGCTAAACAATTTACAATCC	1304
1261	Qy	TAGACAGTAAGCAGTTTTATAAAATACTACAATTTACTCTTTGTGAGTTATGACAGCAGTG	1320
1305	Db	TAGACAGTAAGCAGTTTTATAAAATACTACAATTTACTCTTTGTGAGTTATGACAGCAGTG	1364
1321	Qy	TAACATGTGATAGACATGTAAAGCCCTTTTCAGATTTGTGCAATTTATGAATCTTGATTAATA	1380
1365	Db	TAACATGTGATAGACATGTAAAGCCCTTTTCAGATTTGTGCAATTTATGAATCTTGATTAATA	1424
1381	Qy	TTTTCCTATGCAGATTGCCTCAAAACAGCTTTATATAAAGCACAATTTACTAGTATTTTCACAG	1440
1425	Db	TTTTCCTATGCAGATTGCCTCAAAACAGCTTTATATAAAGCACAATTTACTAGTATTTTCACAG	1484
1441	Qy	TTTTTGTCTAATAGAAAATGCTGATTCGTGATCTCGAGATCAATTTGTGGGAATTTTACATA	1500
1485	Db	TTTTTGTCTAATAGAAAATGCTGATTCGTGATCTCGAGATCAATTTGTGGGAATTTTACATA	1544
1501	Qy	AATCTTTGTTAAATTACTCGAGTGGGCAAGTAGACTTCCTGTCCTTTGCTTTCTTTTTTTTTT	1560
1545	Db	AATCTTTGTTAAATTACTCGAGTGGGCAAGTAGACTTCCTGTCCTTTGCTTTCTTTTTTTTTT	1604
1561	Qy	TCTTTTTGATGCCCTTAATGTAGATATCTTTTATCATCTCTGAATTTGTATTATATATTTAAAG	1620
1605	Db	TCTTTTTGATGCCCTTAATGTAGATATCTTTTATCATCTCTGAATTTGTATTATATATTTAAAG	1664
1621	Qy	TGCTCATTAATAGAAATGATGGATGTAATTTGGATGTAATAATTACTAGTATTAATAATTATA	1680
1665	Db	TGCTCATTAATAGAAATGATGGATGTAATTTGGATGTAATAATTACTAGTATTAATAATTATA	1724
1681	Qy	TCATAATTTGACCCCTGTTGAAATGTGCAATTTATATAAAGCGCAATTTCTTTATCTCT	1739
1725	Db	TCATAATTTGACCCCTGTTGAAATGTGCAATTTATATAAAGCGCAATTTCTTTATCTCT	1783

RESULT 3
US-10-171-581-266
; Sequence 266, Application US/10171581
; Publication No. US2003010426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366

Query Match	48.6%	Score 856.6;	DB 9;	Length 863;
Best Local Similarity	99.5%;	Pred. No. 8.4e-190;		
Matches 859;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
Qy	895	TAGATATTTTCAAAAATACAGTGATCTCATTCGAGGACAATTTTATGGACACACTCACA	954	
Db	1	TAGATATTTTCAAAAATACAGTGATCTCATTCGAGGACAATTTTATGGACACACTCACA	60	
Qy	955	GAGACAGCATTTATGGTCTTTTCAGATAAAAAAGGAAGTC CAGTAAATCTTTGTTGTGG	1014	
Db	61	GAGACAGCATTTATGGTCTTTTCAGATAAAAAAGGAAGTC CAGTAAATCTTTGTTGTGG	120	
Qy	1015	CTCCTGCTGTTTACACCAAGTGAAGAGTCTTTTAGAAAAACAGACCAACAATCTCTGGTATCA	1074	

Db 121 CTCTGCTGTTACACAGTGAAGAGTGTTTTAAAGAAAAACAGACCAAACTCTGGTATCA 180
Qy 1075 GACTGTTTCAGTATGATCTCGTATTAATAATTTATGATATGTTGCGAGTATTAATCTCA 1134
Db 181 GACTGTTTCAGTATGATCTCGTATTAATAATTTATGATATGTTGCGAGTATTAATCTCA 240
Qy 1135 ATCTGACAGAGGCGAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCCTGACCC 1194
Db 241 ATCTGACAGAGGCGAATCTAAAGGGAGAGTCCATCTGGAAGCTGGAGTATATCCTGACCC 300
Qy 1195 AGACCTACGACATGTAAGATTTTCAGCCGGAAGTTTATGATATGATTAAGCAAAATTTA 1254
Db 301 AGACCTACGACATGTAAGATTTTCAGCCGGAAGTTTATGATATGATTAAGCAAAATTTA 360
Qy 1255 CAATCTAGACAGTACGAGTTTATAAATCTACAAATCTCTTTTGTGAGTTATGACA 1314
Db 361 CAATCTAGACAGTACGAGTTTATAAATCTACAAATCTCTTTTGTGAGTTATGACA 420
Qy 1315 GCAGTGTAAATCTGTATAAGACATGTAAGGCCCTTTTCAGATTTTGTGCAATTTATGAATCTTG 1374
Db 421 GCAGTGTAAATCTGTATAAGACATGTAAGGCCCTTTTCAGATTTTGTGCAATTTATGAATCTTG 480
Qy 1375 ATAATATTTCCATGACAGATTCCTCAACAGCTTTTATATAAGCAAAATTAATCTAGTAT 1434
Db 481 ATAATATTTCCATGACAGATTCCTCAACAGCTTTTATATAAGCAAAATTAATCTAGTAT 540
Qy 1435 TCACAGTTTTTCCTATAGAAATGCTGATCTGATTTCTGAGTCAATTTGTGGGAATTT 1494
Db 541 TCACAGTTTTTCCTATAGAAATGCTGATTTCTGAGTCAATTTGTGGGAATTT 600
Qy 1495 TACATAAATCTTTGTTAAATTAATCTAGTGGGCAAGTAGACTTCTGCTTTTGTCTTTT 1554
Db 601 TACATAAATCTTTGTTAAATTAATCTAGTGGGCAAGTAGACTTCTGCTTTTGTCTTTT 660
Qy 1555 TTTTTCCTTTTGTAGCTTAAATGTAGATATCTTTATCATTTCTGAATTTGTATATATAT 1614
Db 661 TTTTTCCTTTTGTAGCTTAAATGTAGATATCTTTATCATTTCTGAATTTGTATATATAT 720
Qy 1615 TTAAGTGTCTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1674
Db 721 TTAAGTGTCTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 1675 ATTATATCTAATTTGACCTTTGTTGAAATTTGTCTTTTATACAAATGAAGCAATTTCTTTA 1734
Db 781 ATTATATCTAATTTGACCTTTGTTGAAATTTGTCTTTTATACAAATGAAGCAATTTCTTTA 840
Qy 1735 TCTCTAAATATGAAAAAATAA 1757
Db 841 TCTCTAAATATGAAAAAATAA 863

RESULT 4

US-10-102-524-990
; Sequence 990, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:

; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; THERAPY AND DIAGNOSIS OF KIDNEY CANCER

; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 415

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-102-524-990

Query Match 23.4%; Score 413.4; DB 9; Length 415;
Best Local Similarity 99.8%; Pred. No. 1.3e-86;
Matches 414; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1202 CGACATTTGAAGATTTTCAGCCGGAAGTTTATGATATGATTAAGCAAAATTTTACAAATCT 1261
Db 1 CGACATTTGAAGATTTTCAGCCGGAAGTTTATGATATGATTAAGCAAAATTTTACAAATCT 60
Qy 1262 AGACAGTAAAGAGTTTATAAATACTACAAATTTCTTTTGTGAGTTATGACAGAGTGT 1321
Db 61 AGACAGTAAAGAGTTTATAAATACTACAAATTTCTTTTGTGAGTTATGACAGAGTGT 120
Qy 1322 AACATGTAAGACATGTAAGGCCCTTTTCAGATTTTGTGCAATTTATGAATTTTACAAAT 1381
Db 121 AACATGTAAGACATGTAAGGCCCTTTTCAGATTTTGTGCAATTTATGAATTTTACAAAT 180
Qy 1382 TTCTATGACAGTTCCTCAACAGCTTTTATATAAGCAAAATTTTACAAATTTTACAAAT 1441
Db 181 TTCTATGACAGTTCCTCAACAGCTTTTATATAAGCAAAATTTTACAAATTTTACAAAT 240
Qy 1442 TTTTCTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTTGTGGGAATTTTACAAAT 1501
Db 241 TTTTCTAATAGAAATGCTGATTTCTGATTTCTGAGATCAATTTTGTGGGAATTTTACAAAT 300
Qy 1502 ATCTTTGTTAAATTAATCTAGTGGGCAAGTAGACTTCTGCTTTTGTCTTTTCTTTT 1561
Db 301 ATCTTTGTTAAATTAATCTAGTGGGCAAGTAGACTTCTGCTTTTGTCTTTTCTTTT 360
Qy 1562 CTTTGTGATGCTTAAATGTAGATATCTTTATCATTTCTGAATTTGTATATATATAT 1616
Db 361 CTTTGTGATGCTTAAATGTAGATATCTTTATCATTTCTGAATTTGTATATATATAT 415

RESULT 5

US-09-815-343-604/c
; Sequence 604, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:

; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.504

; CURRENT APPLICATION NUMBER: US/09/815,343

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 604

; LENGTH: 359

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)-(359)

; OTHER INFORMATION: n = A,T,C or G

US-09-815-343-604

Query Match 14.3%; Score 252.4; DB 10; Length 359;
Best Local Similarity 97.1%; Pred. No. 4.5e-49;
Matches 267; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 1342 AGGCTTTTCAGATTTTGTGCAATTTATGAATTTTCTTATGATTAATTTTCTTATGAGATTCGCTCA 1401
Db 274 AGGCTTTTCAGATTTTGTGCAATTTATGAATTTTCTTATGATTAATTTTCTTATGAGATTCGCTCA 215
Qy 1402 AACAGCTTTTATAAAGCACAAATTTACTAGTATTTTCCACAGTTTTCCTTATGAGATTCGCTCA 1461
Db 214 AACAGCTTTTATAAAGCACAAATTTACTAGTATTTTCCAGCTTTTCTTATGAGATTCGCTCA 155
Qy 1462 GATCTGATTTCTGAGATCAATTTTGTGGGAATTTTATACAAATTTTCTTGTATTAATTTACTGAGT 1521

419	ACCTGAACCTCTCAACAGACACCTGGTTATAAATGTGATCACTAATATGACACCAACCACTCCA	478
421	TGAGAAACTGGGAGAGCGACGTGTACTGGAAATTGTGGAACCGCTTACCAAGCTCATCAG	480
479	GAGTCTCTTTTCCAAATCTCCACAGTCTTTTCCCTCGCTGGGTAAATCATGACTATTGGGCACA	538
481	AGAGGTCTTTTCCAGATATCAAAAGTCTATGTGCTTTGGGAAATCATGATTTTTCACCCCAA	540
539	GGATCAACTCGCTGTAGTCAACAGTAAAGTGTACAAATGTCAGATAGCAAACTCTTGGAAACC	598

541 AAACCAGTTCCTCCAGCTGGGAGTAAACAACATCTACAAATCAGATAGCAGAACCTATGGAAACC 600

Qy	599	ATGGCTAGATGAAGAAGCTATTAGTACTTTAAAGAAAGGTGGTTTTTATTTCACAGAAAGT	658
Db	601	CTGGCTTTAGTAAATGAGTCCATCGCTCTCTTTCAAAAAGGTGCCTTCTACTGTGAGAAGCT	660
Qy	659	TACAACCTAATCCAAACCTTAGG---ATCATCAGTCTAAACACAAACTTGTACTACGGCCC	715
Db	661	GCCGGTCCACGCGGGCTGGGCGAATTGTGTCTCTCAACACCAATCTGTACTATACCAG	720
Qy	716	AAATATAATGACACTGAACAAGACTGACCCAGCCACCAGTTTGAATGGCTACAAAGTAC	775
Db	721	CAATGCGCTGACACAGACATGGCGGACCCTGCCAGCAGTTCCAGTGGCTGGAAGATGT	780
Qy	776	ATTGAACAACTCTCAGCAGAAATAAGGAGAAGGTGTATATCATAGCACATGTTCCAGTGGG	835
Db	781	GCTGACCGATGATCAAAAGCTGGGGACATGTTGTATCATTTGCGGCCACGTGCCCCCGGG	840
Qy	836	GTATCTGCCATCTTCACAGAACATCA CAGCAATGAGAGAAATACTATAATGAGAAATTTGAT	895
Db	841	GTCTTTTGAGAGACGCAAAA CAAAGCATGGTTCGGGAGGGCTTCAATGAAAAATACCT	900
Qy	896	AGATATTTTCAAAAATA CAGTGTATGATTCAGAGACAATTTTTATGGACACACTCACAG	955
Db	901	GAAGGTGGTCCGGAAGCATCATCGCTCATAGCAGGGCAGTTCTTCGGGCACCACACAC	960
Qy	956	AGACAGCATATGGTCTTTTCAGATAAAAAGGAAGTCCAGTAAATTTCTTTGTTTGTGGC	1015
Db	961	CGACAGCTTTTCGGATGCTCTATGATGATGAGGTGCCCATGAAGGCCCATGTTTCATCAC	1020
Qy	1016	TCTGCTGTTTACACCAAGTCAAGA-----GTGTTTTAGAAAAACAGACCAACAATCC	1066
Db	1021	ACCTGGAGTCACCCCATGAAAACCACTTACCTGGAGTGGTCAATGGGGCCCAACAATCC	1080
Qy	1067	TGTTATCAGACTGTTTTCAGTATGATCCTCGTGATTATAAAATTATTCGATGTGTGCAGTA	1126
Db	1081	AGCCATCCGGGTTCGAATATGACCGAGCCACACTGAGCCTNNAGGACATGGTGACCTA	1140
Qy	1127	TTACTTTGAATCTGACAGAGCGGCAATCTAAAGGGAGAGTCCATCTCTGGAAGCTGGAGTATAT	1186
Db	1141	CTTCATGAACCTTGAGCCAGCGGCAATGCTCTAGGSGGACGCCGCGCTGGGAGCTCGAGTACCA	1200

```

1187 CCTGACCCGACCTACG 1203
      |||||
      |||||
Db      1201 GCTGACCGAGGCGCTATG 1217

RESULT 7
US-09-967-768A-319
; Sequence 219, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111

```

RESULTS

```

US-09-967-768A-219
; Sequence 219, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111

```

;; PRIOR FILING DATE: 2000-09-28
;; NUMBER OF SEQ ID NOS: 325
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 219
;; LENGTH: 1610
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(1610)
;; OTHER INFORMATION: n=a,t,g or c
US-09-967-768A-219

Query Match 13.8%; Score 243.8; DB 10; Length 1610;
Best Local Similarity 53.6%; Pred. No. 1.2e-46; Indels 12; Gaps 2;
Matches 556; Conservative 0; Mismatches 469;

QY 179 AGGACAGTTTGGCATGTGACTTACACTTAGACCCCTACTTACACATCACAGATGA 238
DB 181 AGGGAAGTTCTGGCACATCGCTGACCTGCACCTTGACCCCTGACTACAAAGTATCCAAAGA 240
QY 239 CCACACAAAAGTGTGCTTCATCTAAAGTGCAATGCCTCAACCCCTGGCCCTTTGG 298
DB 241 CCCCTTCCAGGTGTGCCCATCAGCTGGATCCCGAGCGTGCCTGGAGCGCCCTGGGG 300
QY 299 AGATGTTCTGTGATTCTCCATATCAACTTATTTTGTGACGATTTGATTTTATAAAA 358
DB 301 TGACTACTCTGTGATTTCTCCCTGGGCCCTCACTCACTCTCCATCTATGCCATGAAGA 360
QY 359 TTCTGGACAAAGACATCTTTTCATGATATGGACAGGGGATAGCCACCTCATCTTCTGT 418
DB 361 GATTGAGCCAGAGCCAGACTTCTCTGCGCTGGTGTGATGACACGCTCATGTGCCCGA 420
QY 419 ACCTGAACTCTCAACAGACACTGTTTAAATGTGATCACTAATATGACAAACCATCCA 478
DB 421 TGAGAAACTGGGAGAGCGAGCTGTACTTGGAAATTTGTGAAACGCTGACCAAGCTCATAG 480
QY 479 GAGTCTCTTTCCAAATCTCCAGGTTTTCCTCGCTGGGTAAATCATGACTATTGGCCACA 538
DB 481 AGAGGCTTTCCAGATACTAAGTCTATGCTGTTTGGGAAATCATGATTTTCCACCCAA 540
QY 539 GGATCAACTGCCTGTAGTCAACAGTAAAGTGTACAATGACAGTAGCAAAACCTCTGGAACC 598
DB 541 AAACAGTTCCTCCAGCTGGAAGTAAACATCTACAATCAGATAGCAGAACTATGGAACC 600
QY 599 ATGGCTAGATGAAGAAGCTATTAGTACTTTAAGAAAGGTGTTTTTATTCAGAAAGT 658
DB 601 CTGGCTTAGTAATGAGTCCATCGCTCTCTTCAAAAAGGTGCTTCTACTGTGAGAAGCT 660
QY 659 TACAACCTTAATCCAAACCTTAGG--ATCATCAGTCTAAACACAACTTGTACTACGGCCC 715
DB 661 GCGGGTCCAGCGGGGCTGGGGAATTTGGTCTCTCAACACCAATCTGTACTATACCAG 720
QY 716 AAATATAATGACACTGAACAAGACTGACCCAGCCAAACAGTTTGAATGGCTAGAAAGTAC 775
DB 721 CAATGGCTGACAGCAGACATGCGGACCTGCGGACAGTTCAGTGGCTGGAAGATG 780
QY 776 ATTGAACACTCTCAGCAGAAATAGGAGAAGGTGTATATCATAGCAGATGTTCCAGTGG 835
DB 781 GCTGACCGATGTCATCCAAAGCTGGGACATGGTGTACATTTGTGCGGCACGCTGCCCGGG 840
QY 836 GTATCTGCCATCTTTCACAGAACTACAGCAATGAGAGAACTATATAATGAGAAATTTGAT 895
DB 841 GTTCTTTGAGAAACGCAAAACAGGATGTTTCCGGAGGGCTTCAATGAAATAATACCT 900
QY 896 AGATATTTTTCAAAATACAGTATGATGATTCAGGAGCAATTTTATGGACACACTCACAG 955
DB 901 GAAGGTGGTCCGAAGCATCATCGCTCATAGCAGGCGAGTCTTCTGGGCAACCAACACAC 960
QY 956 AGACAGCAATTATGTTCTTTCAGATAAAAGAGGAGTCCAGTAAATCTTTGTTTGGC 1015
DB 961 CGACAGCTTTTCGATGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020

QY 1016 TCCTGCTGTTACACCAAGTGAAGA-----GTGTTTAAAAACAGACCAACAATCC 1066
DB 1021 ACCTGGAGTCAACCCATGGAACACACATTACCTGGAGTGGTCAATGGGCGCAACAATCC 1080
QY 1067 TGGTATCAGACTGTTTTCAGATGATGCTCTCGTGAATTAATAATTATTGGATATGTTGAGTA 1126
DB 1081 AGCCATCCGGGTGTTTGAATATGACCGAGCCACACTGAGCCCTNNAGGACATGTTGACCTA 1140
QY 1127 TTACTTGAATCTGACAGAGCGCAATCTAAAGGAGAGTCCATCTGGAAGCTGAGTATAT 1186
DB 1141 CTTTCATGAACCTGAGCCAGCGGAATGCTCAGGGGACGCGCTGGAGCTCGAGTACCA 1200
QY 1187 CTTGACCCAGACCTACG 1203
DB 1201 GCTGACCGAGGCTATG 1217

RESULT 8

US-09-815-343-1053/c
; Sequence 1053, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1053
; LENGTH: 156
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(156)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-1053

Query Match 8.2%; Score 144.4; DB 10; Length 156;

Best Local Similarity 92.9%; Pred. No. 4.2e-24;
Matches 145; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1461 TGATCTGATTCGAGATCAATTTGTGGGAATTTTACATAAATCTTTGTTAATTAAGT 1520
DB 156 TGATTTGATTCGANANCAATTTGTGGGAATTTTACANAAATNTTTGTTAANTACTGAG 97
QY 1521 TGGGCAAGTAGACTTCCTGCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1580
DB 96 NCGGCAAGTANACTTCCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1616
QY 1581 AGATATCTTTATCATCTCGAATTTGTTATATATATTT 1
DB 36 AGATATCTTTATCATCTCGAATTTGTTATATATATTT 1

RESULT 9

US-09-783-590-5535
; Sequence 5535, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15

PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5535
LENGTH: 142
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (44)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (96)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (130)
OTHER INFORMATION: n equals a,t,g, or c
US-09-793-590-5535

Query Match 6.3%; Score 110.8; DB 10; Length 142;
Best Local Similarity 92.0%; Pred. No. 2.7e-16;
Matches 126; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

Qy 25 GCTCGGGACAGCCCGAACCTCCAGGTCCAGCCCGCGCCCTCCATGGCGCTGGTGGCG 84
Db 7 GCACAGGACAGCCCGAACCTCCAGGTCCAGCCCGCGCCCTCCATGGCGCTGGTGGCG 66

Qy 85 CACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 144
Db 67 CACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126

Qy 145 GCCTCGGACAGCGCGGCGAGG 161
Db 127 -GCCNCTAGCGGGGCGAG 142

RESULT 10
US-10-066-543-2024/c
Sequence 2024, Application US/10066543
Publication No. US20030087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqiu
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2024
LENGTH: 381
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-2024

Query Match 4.4%; Score 77.8; DB 9; Length 381;
Best Local Similarity 52.5%; Pred. No. 2.5e-08;
Matches 200; Conservative 0; Mismatches 172; Indels 9; Gaps 1;

Qy 811 ATATCATGACATGTTCCAGTGGGTATCTGCCATCTTCACAGAACATCACAGCAATGA 870

Db 381 ACATTGTCGCCACGTGCCCCGGGTCTTTTGAGAGACGCAAAACAAGGCATGGTTCC 322
Qy 871 GAGAACTACTAATAAGAAATGATAGATATTTTCAAAAATACAGTGTGTCATTGTCAG 930
Db 321 GGGAGGGCTTCAATGAAAAATACCTGAAGTGGCCCGGAAGCATCATCGCGTCATAGCAG 262
Qy 931 GACAATTTTATGGACACACTCCACAGACAGCAGCATTTATGTTCTTTTCAGATAAAAAAGGAA 990
Db 261 GGCAGTCTTCGGGGCACCAACACACCGCAGAGCTTTCCGATGCTTATGATGATGCGCGTG 202
Qy 991 GTCCAGTAAATCTTTTGTGCTCTCTGCTTACACACAGTGAAGA-----GTG 1041
Db 201 TCCCATAGCGCCATGTTTCATCACACCTGGAGTCAACCCATGGAACACCATTTACCTG 142
Qy 1042 TTTTAAAAAACAAGCAACCAATCTGGTATCAGACTGTTTTCAGTATGATCTCTGATGATT 1101
Db 141 GAGTGTGTCATGGGGCAACAATCCAGCCATCCGGGTGTTTCAATATATGACCGAGCCACAC 82
Qy 1102 ATAAATTTATGGATATGTTGAGTATTTACTTGAATCTGACAGAGGGGAATCTTAAGGGAG 1161
Db 81 TGAGCCTGAAGGACATGGTGACCTTCTCATGAACCTGAGCCAGGCGAATGCTCAGGGGA 22
Qy 1162 AGTCCATCTGGAAGCTGGAGT 1182
Db 21 CGCGCGCTGGGAGCTCGAGT 1

RESULT 11
US-10-066-543-1700
Sequence 1700, Application US/10066543
Publication No. US20030087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqiu
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1700
LENGTH: 346
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-1700

Query Match 3.7%; Score 65.6; DB 9; Length 346;
Best Local Similarity 94.4%; Pred. No. 1.7e-05;
Matches 68; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1691 ACCCTTGTGAAATGTCATTTATACAAATGAAGCGAATTTCTTTATCTTAATATGAAA 1750
Db 1 ACCCTTGTGAAATGTCATTTATACAAATGAAGCGAATTTCTTTATCTTAATATGAAA 60

Qy 1751 AAAAAA 1762
Db 61 AAAAAA 72

RESULT 12
US-10-239-676-6
Sequence 6, Application US/10239676

```
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 6
; LENGTH: 7544
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5455)
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-239-676-6

Query Match 3.5%; Score 61.4; DB 9; Length 7544;
Best Local Similarity 51.0%; Pred. No. 0.0011;
Matches 176; Conservative 0; Mismatches 161; Indels 8; Gaps 1;

QY 1413 ATAAACACAACTACTAGTATTTACAGTTTTTGTCTAAATAGAAAATGCTGATTCGATTC 1472
Db 42 AAAAATAAATTAATAAATAAGAAATATAATAATTTTAAATAATTTTCGGAAT 101
QY 1473 TGAGATCAATTTCTGGGAATTTTACATAATCTTTGTTAAATCTGAGTGGGCAAGTAGA 1532
Db 102 GGAATAATTTTAAATTTGATATAAATTTAGGCTTTATAAGGAAGGATAAGTAAA 161
QY 1533 CTTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1592
Db 162 -----TTTTTTTTTTTTTTTTTTTTTTTTTATATATATATATATATATATATATAT 213
QY 1593 CATTCTGAATTTGATATATATATATATATATATATATATATATATATATATATATAT 1652
Db 214 TATTTCGATATAGGAAAAAATAATAATATAAATTTGGAATAAATAATAAGTAATTTT 273
QY 1653 ATGTAATATTCAGTTTATATATATATATATATATATATATATATATATATATATATAT 1712
Db 274 ATTAAGTAATTCGGTTTAAATTTTAAATAATAAGTTTATATATATATATATATATATAT 333
QY 1713 ATCAATAAAGCGAATCTTTTATCTCTAAATATGAAAAAATAA 1757
Db 334 TAAAGAAAAATGTAAGGATATGAATGTTTATAGTATAGTATAGTATAGTATAGTATAGT 378

RESULT 13
US-10-239-676-164
; Sequence 164, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7

; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: EpiGenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/172,086
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 77
; LENGTH: 4858
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-172-086-77

Query Match 3.1%; Score 55.4; DB 9; Length 4858;
Best Local Similarity 47.8%; Pred. No. 0.021;
Matches 161; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 1424 TTACTAGTATTTACAGTTTTTGTCTAAATAGAAAATGCTGATTCGATTCGATCAAT 1483
Db 1424 TTACTAGTATTTACAGTTTTTGTCTAAATAGAAAATGCTGATTCGATTCGATCAAT 1483
```


[illegible]

RESULT 15

```

US-10-198-846-1369/c
/ Sequence 1369, Application US/10198846
/ Publication No. US2003009974A1
/ GENERAL INFORMATION:
/ APPLICANT: Lillie, James
/ APPLICANT: Xu, Yongyao
/ APPLICANT: Wang, Youzhen
/ APPLICANT: Steinmann, Kathleen
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ TITLE OF INVENTION: THERAPY OF BREAST CANCER
/ FILE REFERENCE: WFI-049
/ CURRENT APPLICATION NUMBER: US/10/198,846
/ CURRENT FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/306,220
/ PRIOR FILING DATE: 2001-07-18
/ NUMBER OF SEQ ID NOS: 14084
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1369
/ LENGTH: 539
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ NAME/KEY: misc feature
/ FEATURE:
/ LOCATION: 2, 3, 4, 8, 14, 25, 113, 117, 118, 121, 123, 125, 127, 129,
/ LOCATION: 130, 134, 135, 142, 145, 152, 155, 156, 164, 167, 169, 189,
/ LOCATION: 190, 201, 202, 211, 238, 239, 270, 271, 292, 295, 302, 303,
/ LOCATION: 306, 314, 315, 321, 339, 346, 349, 359, 374, 388, 395
/ OTHER INFORMATION: n = A,T,C or G
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 396, 410, 414, 430, 432, 435, 439, 440, 442, 443, 447, 448,
/ LOCATION: 453, 473, 476, 480, 482, 486, 487, 493, 495, 500, 501, 510,
/ LOCATION: 530
/ OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1369

```

	Query Match	3.18	Score 55.2	DB 9	Length 539
	Best Local Similarity	47.28	Pred. No. 0.0059		
	Matches 108	Conservative 0	Mismatches 121	Indels 0	Gaps 0
OY	1534	TTCTGTCTTGGTCTCTTTTTTTTTTTCTTTTTGATGCCTTAATGTAGATATCTTTATC	1593		
Db	308	TTTATNNATTTTCANTTTTTTTTTTTTTTTTTTTTTTANNTTTTTTTTTATTTATTTTTTTT	249		
OY	1594	ATTCCTGAATGTATATATATATTTAAAGTCTCATTAATAGAATGATGAATGAATTTGGA	1653		
Db	248	TTTTTTTTTTNNTTNNATTTTTTTTTNN	189		

[illegible]

Search completed: June 28, 2003, 12:28:47
Job time : 169.577 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:12:40 ; Search time 1630.98 Seconds
(without alignments)
17516.326 Million cell updates/sec

Title: US-09-981-353-47

Perfect score: 1764

Sequence: 1 ctccaggctcgaagctccgag.....tgaaaaaaggg 1764

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	936.8	53.1	1717	11 AK002455	AK002455 Mus muscu
2	875.4	49.6	919	9 AL513818	AL513818 AL513818
3	733.6	41.6	810	13 BI820750	BI820750 603034346
4	725.4	41.1	837	9 AI379164	AI379164 tc66b07.x
5	714.2	40.5	815	13 BI762966	BI762966 603048239
6	692	39.2	954	12 BG674083	BG674083 602620079

7	676.2	38.3	816	13 BI757489	BI757489 603029636
8	656.8	37.2	726	9 AA873600	AA873600 obl2a02.8
9	648.2	36.7	809	12 BG567159	BG567159 602589619
10	642	36.4	739	14 BM985118	BM985118 UI-CF-EC1
11	635	36.0	635	14 BM795153	BM795153 K-EST0076
12	631.6	35.8	650	13 BI759658	BI759658 603045507
13	617.2	35.0	663	12 BG432589	BG432589 602500747
14	611.2	34.6	682	10 AW978459	AW978459 EST300568
15	596.2	33.8	602	10 BE222616	BE222616 hu49a09.x
16	588.4	33.4	748	13 BI756218	BI756218 603024250
17	582.8	33.0	873	12 BG399988	BG399988 602442042
18	565.4	32.1	766	13 BI249553	BI249553 602996209
19	563.4	31.9	566	13 BI962992	BI962992 ie61d03.Y
20	551.8	31.3	567	9 AI493580	AI493580 th39a09.x
21	545.6	30.9	600	12 BE930320	BE930320 RC4-GN006
22	544.4	30.9	546	10 BE280512	BE280512 601155383
23	544.2	30.9	794	9 AU080158	AU080158 AU080158
24	543.8	30.8	567	9 AI672280	AI672280 wg31c05.x
25	531.6	30.1	570	9 AI769126	AI769126 wg34b01.x
26	529.4	30.0	818	12 BG619567	BG619567 602618962
27	528.4	30.0	1272	13 BI554723	BI554723 603236584
28	527	29.9	1101	14 BM805174	BM805174 AGENCOURT
29	524.6	29.7	785	12 BG169392	BG169392 602320958
30	520.2	29.5	742	13 BI654512	BI654512 603281094
31	520	29.5	779	14 BQ042425	BQ042425 UI-M-EMO-
32	517.4	29.3	527	12 BF114920	BF114920 hr70b10.x
33	513	29.1	876	13 BI645836	BI645836 603275501
34	510.6	28.9	785	13 BI647402	BI647402 603279663
35	504.2	28.6	927	14 BQ715296	BQ715296 AGENCOURT
36	500.4	28.4	581	12 BG538467	BG538467 602567175
37	500.4	28.4	715	13 BI557227	BI557227 603238824
38	489	27.7	500	9 AI005229	AI005229 oc98f07.x
39	486	27.6	486	9 AI285540	AI285540 qu49a04.x
40	485.4	27.5	751	10 AW475628	AW475628 un69c11.Y
41	483.8	27.4	543	10 AV746635	AV746635 AV746635
42	481.2	27.3	494	9 AI222102	AI222102 gg96d04.x
43	480	27.2	499	14 N31113	N31113 YX52c01.r1
44	475.6	27.0	754	9 AI529785	AI529785 ui82d11.Y
45	475	26.9	493	9 AA830466	AA830466 oc51f10.8

ALIGNMENTS

RESULT 1	AK002455	AK002455	1717 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK002455	Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610010C24:acid sphingomyelinase-like phosphodiesterase 3a, full insert sequence.				
ACCESSION	AK002455	AK002455.1	GI:12832451			
VERSION	AK002455	HTC; CAP trapper.				
KEYWORDS	AK002455	Mus musculus (strain:CS7BL/6J) adult male kidney cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library				
SOURCE	AK002455	clone:0610010C24.				
ORGANISM	AK002455	Mus musculus				
REFERENCE	AK002455	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	AK002455	1 Carninci, P. and Hayashizaki, Y.				
TITLE	AK002455	High-efficiency full-length cDNA cloning				
JOURNAL	AK002455	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	AK002455	99279253				
PUBMED	AK002455	10349636				
AUTHORS	AK002455	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	AK002455	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	AK002455	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	AK002455	20499374				
PUBMED	AK002455	11042159				

[illegible]

QY 940 ATGCACACACTCTACAGAGACAGCATTTATGTTCTTTTCAGAT--AAAAAGGAGTCCAGT 997
DB 421 ATGCACACACTCTACAGAGACAGCATTTATGTTCTTTTCAGATCAACAAGGAGTCCAGT 480
QY 998 AATATCTTTGTTGTTGCTCTGCTGTATACACAGTGAAGAGTGTGTTTGAAGAAACAGAC 1057
DB 481 AATATCTTTGTTGTTGCTCTGCTGTATACACAGTGAAGAGTGTGTTTGAAGAAACAGAC 539
QY 1058 CAACAATCTGTTATCAGACTGTTTCAGTATGATCTCTGATATTAATAATTTATGGATAT 1117
DB 540 CAACAATCTGTTATCAGACTGTTTCAGTATGATCTCTGATATTAATAATTTATGGATAT 599
QY 1118 GTTCAGTATTAATCTGATCTGACAGAGCGGAATCTAAAGGAGAGTCCATCTGG-AAGC 1176
DB 600 GTTCAGTATTAATCTGATCTGACAGAGCGGAATCTAAAGGAGAGTCCATCTGGAAGC 659
QY 1177 TGGAGTATATCTGACACAGACTTAC-CACATTTGAGATTTTGCAGCGGAAAGTTTATAT 1235
DB 660 TGGAGTATATCTGACACAGACTTACCGACATTTGAAGATTTTGCAGCGGAAAGTTTATAT 719
QY 1236 GGATTAGTAAACAATTTACAA--TCCTAGACAGTAAGCAGTATTAATAATACTACAAATT 1293
DB 720 GATTAGTAAACAATTTACAAATCTTAGAACAGTAAGCAGTATTAATAATACTACAAATT 779
QY 1294 ACTTCTTTGCTGATGATGA-CAGCAGTGTAAATGTTGATGAAGCATGTAAGGCTTTTCAG 1352
DB 780 CTCTTCTGCTGATGATGACAGCAGTGTATCATGTGATTAGGACATGTAGGCGCTTCCG 839
QY 1353 ATTTGTGCAATTTATGAATCTTTGATAATTTTCTTATGCAATTTGCTTCAACAGCTTTAT 1412
DB 840 ATTTGTGCTTATGACATCTGATCTA-TTCTATGCAATGCTTTTCAAGAGTTATAT 898
QY 1413 ATAAGACCAATTAAT 1428
DB 899 TAGCCCAATTAATAGT 914

RESULT 7
BI757489
LOCUS 816 bp mRNA linear EST 25-SEP-2001
DEFINITION 603029636f1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5199800 5',
mRNA sequence.
ACCESSION BI757489
VERSION BI757489.1 GI:15749067
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11500 row: h column: 09
High quality sequence stop: 783.
Location/Qualifiers
1..816
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5199800"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;

FEATURES
source

Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH MGC Library."

BASE COUNT 207 a 229 c 178 g 202 t
ORIGIN

Query Match 38.3%; Score 676.2; DB 13; Length 816;
Best Local Similarity 98.5%; Pred. No. 1.3e-105;
Matches 714; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

QY 1 CTCAGCGCTGACGGTCCGAGTGGAGCTGGGGACAGCCGACCTCCAGGTCAGCCCGC 60
DB 37 CTCAGCGCTGACGGTCCGAGTGGAGCTGGGGACAGCCGACCTCCAGGTCAGCCCGC 96
QY 61 GGGCTTCCATGCGGCTGGTGGCGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 97 GGGCTTCCATGCGGCTGGTGGCGCACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 156
QY 121 GCTCGGGCTCGGGCTGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 157 GCTCGGGCTCGGGCTGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 216
QY 181 GACAGTTTGGCATGTGACTGACTTACACTTTAGACCTTACTTACACATCACAGATGACC 240
DB 217 GACAGTTTGGCATGTGACTGACTTACACTTTAGACCTTACTTACACATCACAGATGACC 276
QY 241 ACACAAAAGTGTGCTTTCATCTAAAGTGCAAAATGCTTCCAACTGCGGCTTTGGAG 300
DB 277 ACACAAAAGTGTGCTTTCATCTAAAGTGCAAAATGCTTCCAACTGCGGCTTTGGAG 336
QY 301 ATGTTCTGTGTGATTTCTCCATATCAACTTATTTTGTGAGCATTTGATTTTATTAATAAT 360
DB 337 ATGTTCTGTGTGATTTCTCCATATCAACTTATTTTGTGAGCATTTGATTTTATTAATAAT 396
QY 361 CTGGACAAGAGCATCTTTTCAATGATGAGAGGGATAGCCCACTCATGTTCTCTGTAC 420
DB 397 CTGGACAAGAGCATCTTTTCAATGATGAGAGGGATAGCCCACTCATGTTCTCTGTAC 456
QY 421 CTGAACCTCTCAACAGACACTGTTTATTAATGATCACTAATATGACACACCATTCAGA 480
DB 457 CTGAACCTCTCAACAGACACTGTTTATTAATGATCACTAATATGACACACCATTCAGA 516
QY 481 GTCTCTTTTCCAAATCTCCAGGTTTTCCTCGCTGGGTATCATGACTATTTGGCCACAGG 540
DB 517 GTCTCTTTTCCAAATCTCCAGGTTTTCCTCGCTGGGTATCATGACTATTTGGCCACAGG 576
QY 541 ATCAACTGCTGTAGTACCAAGTAAAGTGTACAAATGATGAGTAAACCTCTGGAACCAT 600
DB 577 ATCAACTGCTGTAGTACCAAGTAAAGTGTACAAATGATGAGTAAACCTCTGGAACCAT 636
QY 601 GGCTAGTAGAAGAGCTATTAGTAC-TTTAAGGAAAGGTGGTGTATTTTATTCACAGAAG-T 658
DB 637 GGCTAGTAGAAGAGCTATTAGTAC-TTTAAGGAAAGGTGGTGTATTTTATTCACAGAAG-T 696
QY 659 TACAACCTTAACCAACCTTAGGATCATC-AGTCTAAACACAACTTTGTACTAGGCCCAA 717
DB 697 TACAACCTTAACCAACCTTAGGATCATC-AGTCTAAACACAACTTTGTACTAGGCCCAA 756
QY 718 ATATA 722
DB 757 TATTA 761

RESULT 8
AA873600/c
LOCUS
DEFINITION

PHOSPHODIESTERASE ;, mRNA sequence.

ACCESSION AA873600
 VERSION AA873600.1 GI:2969722
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 726)
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 National Cancer Institute, NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 366.
 Location/Qualifiers
 1..726
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1323434"
 /clone_lib="NCI CGAP_Kid3"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p773D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer,
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified p773 vector. mRNA
 source: 2 pooled kidneys. Library went through one round
 of normalization. Library constructed by Bento Soares and
 M. Fatima Bonaldo."

BASE COUNT 267 a 129 c 117 g 213 t
 ORIGIN

Query Match 37.2%; Score 656.8; DB 9; Length 726;
 Best Local Similarity 96.7%; Pred. No. 2.9e-102;
 Matches 691; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 1036 AGAGTGTGTTTGAAGAAACAGACCAACAATCCTGGTATCAGACTGTTTCAGTATGATCCCTC 1095
 DB 704 AGGTTTGTGAAACCCGTCACCATCCCTGTTATCAGACTGTTTCGATGATCCCG 645

QY 1096 GTGATTATAAATATTGGATATTGTCAGTATTACTTGAATCTGACAGAGCGCAATCTAA 1155
 DB 644 GTGATTATAAATATTGGGATATTGTCAGTATTACTTGAATCTGACAGTGGCGAATCTAA 585

QY 1156 AGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACACAGACCTACGACATTTGAGATT 1215
 DB 584 AGGAGAGTCCATCTGGAAGCTGGAGTATATCTGACACAGACCTACGACATTTGAGATT 525

QY 1216 TGCAGCCGGAAGTTTATATGAGTATAGTAAACAAATTTACAATCTCAGACAGTAAGCAGT 1275
 DB 524 GTACGCCGAAGTTTATATGAGTATAGTAAACAAATTTACAATCTCAGACAGTAAGCAGT 465

QY 1276 TTATAAATACTACAATTAATCTTTGTGAGTTATGACAGAGTGTAAATGTGTAAGA 1335
 DB 464 TTATAAATACTACAATTAATCTTTGTGAGTTATGACAGAGTGTAAATGTGTAAGA 405

QY 1336 CATGTAAAGCCCTTCAGATTGTGCAATTTAGATCTTCATATATATTTTCTTCCATGAGATT 1395
 DB 404 CATGTAAAGCCCTTCAGATTGTGCAATTTAGATCTTCATATATATTTTCTTCCATGAGATT 345

1396 GCCTCAACAGCTTTTATATAAAGCAACAATTAAGTATTTACAGTATTTTGCTAATAGAA 1455

344 GCCTCAACAGCTTTTATATAAAGCAACAATTAAGTATTTACAGTATTTTGCATAAGAA 285
 QY 1456 AATGCTGATTCGATTCGAGATCAATTTTGGGAATTTTACATAAAATCTTTGTTAATA 1515
 DB 284 AATGCTGATTCGATTCGAGATCAATTTTGGGAATTTTACATAAAATCTTTGTTAATA 225

QY 1516 CTGAGTGGGCAAGTAGACTTCCTGCTCTTCTTC-TTTTTTTTTTCTTTTGTATGCT 1574
 DB 224 CTGAGTGGGCAAGTAGACTTCCTGCTCTTCTTC-TTTTTTTTTTCTTTTGTATGCT 165

QY 1575 TAATGTAGATATCTTTATCATCTCTGAATGTTATATATATATTAATAAGTCTCATTAATA 1634
 DB 164 TAATGTAGATATCTTTATCATCTCTGAATGTTATATATATTAATAAGTCTCATTAATA 105

QY 1635 ATGATGGATGAAATTTGGATGTAATATATTCAGTTTATATATATATTAATTTGTACCC 1694
 DB 104 ATGATGGATGAAATTTGGATGTAATATATTCAGTTTATATATATATTAATTTGTACCC 45

QY 1695 TTGTTGAATTTGTCATTTTATACAATAAAGCAATCTTTTATCTC 1738
 DB 44 TTGTTGAATTTGTCATTTTATACAATAAAGCAATCTTTTATCTC 1

RESULT 9
 LOCUS BG567159
 DEFINITION 602589619F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4723589 5',
 mRNA sequence.
 ACCESSION BG567159
 VERSION BG567159.1 GI:13574812
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 809)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1583 row: f column: 06
 High quality sequence stop: 629.
 Location/Qualifiers
 1..809
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4723589"
 /clone_lib="NIH MGC 76"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggcccttcggcc); Site 2: SfiI (ggccattggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGCCATATATGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGCGCGCATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.85
 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH MGC Library."

BASE COUNT 260 a 121 c 162 g 264 t
 ORIGIN

Query Match 36.7%; Score 648.2; DB 12; Length 809;
 Best Local Similarity 91.0%; Pred. No. 7.8e-101;

	Matches	722;	Conservative	0;	Mismatches	65;	Indels	6;	Gaps	3;
QY	873	GAATACTACTAAATGACAGAAATTGATAGATATTTTCCAAAATACAGTGATGTCAITGCAGGA	932							
Db	1	GAATACTACTAAATGACAGAAATCGTAGATATTTTTCAAATAACAGTGATGTCAITGCAGGA	60							
QY	933	CAATTTTATGGACACACTCAGACAGACAGCATTAATGGTTCTTTTCAGATAAAAAAGGAAGT	992							
Db	61	CNA'TTTATGGACACACTCAGACAGACAGCATTAATGGTTCTTTTCAGATAAAAAAGGAAGT	120							
QY	993	CCAGTAAATTCCTTTGTTGGTGGCTCCTGCTGTTACACACAGTGAAGAGTGTTTTAGAAAAA	1052							
Db	121	CCAGTAAATTCCTTTGTTGGTGGCTCCTGCTGTTACACACAGTGAAGAGTGTTTTAGAAAAA	180							
QY	1053	CAGACCACAACATCCTGCTGATCAGACTGTTTCAGTATGATCCTCGTGATTATAAATATTATG	1112							
Db	181	CAGACCACAACATCCTGCTGATCAGACTGTTTCAGTATGATCCTCGTGATTATAAATATTATG	240							
QY	1113	GATATGTTGCGATTAATCTTGAATCTGACAGAGCGCAATCTAAAGGAGNGTCCATCTCG	1172							
Db	241	GATATGTTGCGATTAATCTTGAATCTGACAGAGCGCAATCTAAAGGAGNGTCCATCTCG	300							
QY	1173	AAGCTGGAGTATATCCTGACCCAGACTCAGCAATTTGAAGAATTTGCAGCGGAAAAGTTTA	1232							
Db	301	AAGCTGGAGTATATCCTGACCCAGACTCAGCAATTTGAAGAATTTGCAGCGGAAAAGTTTA	360							
QY	1233	TATGGATTAGCTAAAACAATTTTACAATCCTAGACAGTAAGCAGTTTATAAATACTACAA	1292							
Db	361	TATGGATTAGCTAAAACAATTTTACAATCCTAGACAGTAAGCAGTTTATAAATACTACAA	420							
QY	1293	TACTTCTTTGTGAGTTATCAGACAGCAGTGTAACATGTGATGAACATGTAAGGCCCTTTCCAG	1352							
Db	421	TACTTCTTTGTGAGTTATCAGACAGCAGTGTAACATGTGATGAACATGTAAGGCCCTTTCCAG	480							
QY	1353	ATTTCGTGCAATTTATGAATCTTGATTAATATTTCTTATGCAGATTTGCTCAAAACAGCTTTAT	1412							
Db	481	ATTTCGTGCAATTTATGAATCTTGATTAATATTTCTTATGCAGATTTGCTCAAAACAGCTTTAT	539							
QY	1413	ATAAAGCACAAATTAAGTATTTACAGTTTTTCAGTTTTTCTTAATAGAAAAATGCTGATTC	1472							
Db	540	ATAAAGCACAAATTAAGTATTTACAGTTTTTCTTAATAGAAAAATGCTGATTC	599							
QY	1473	TGAGATCAATTTGTGGGAATTTTACATAAATCTTTGTTAAATCTAGTGGGCAAGT--A	1530							
Db	600	TGAGATCAATTTGTGGGAATTTTACATAAATCTTTGTTAAATCTAGTGGGCAAGT--A	659							
QY	1531	GACTTCCTGCTTTTGGCTTT---CTTTTTTTTTTCTTTTGTATGTCCTTAATGATGATATC	1587							
Db	660	AAATTCCTGCTTTTGGCTTTCCCATATATATAGTCCTTTTTGGAGAGACTTAATGTGTATATC	719							
QY	1588	TTTATCAATTCGAATTCGATATATATTTAAAGTGCTCAATTAATAGAATGATGGATGTA	1647							
Db	720	TTTATCACTCTGAATGGGTTTATTTTTTTAAAGGGCTTTAATAGAAAGTGGGCGGTAA	779							
QY	1648	ATTGGATGTAAAT	1660							
Db	780	CTGCGGTTGAAAT	792							

[illegible]

Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Tel: 319 356 4866
Fax: 319 356 5671
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA-yes

FEATURES
SOURCE

```

FEATURES
source
POLYA=yes.
1. .739
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-ace-o-19-o-UI"
/clone_lib="UI-CF-EC1"
/tissue_type="Lung"
/dev_stages="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (Tl phage resistant)"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGCCTTAC.
TAG LIB=UI-CF-EC1
TAG TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG SEQ=AAGTGCCTTAC"
264 a 123 c 103 g 249 t
BASE COUNT
ORIGIN

```

	Query Match	36.4%;	Score 642;	DB 14;	Length 739;
	Best Local Similarity	98.5%;	Pred. No. 9.3e-100;		
	Matches 648;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;
QY	1100	TTATAATTAATTCGATATGTTGCAGTATTACTTGAATCTGCAGAGGCGAATCTTAAGGG	1159		
Db	658	TCAGATAAAAAAGGATATGTTGCAGTATTACTTGAATCTGCAGAGGCGAATCTTAAGGG	599		
QY	1160	AGAGTCCATCTGGAAGCTGGAGTATATCTGTGCCAGACCTACGACATTGAAGATTTCGA	1219		
Db	598	AGAGTCCATCTGGAAGCTGGAGTATATCTGTGCCAGACCTACGACATTGAAGATTTCGA	539		
QY	1220	GCCGGAAGAGTTTATATGGATTAGCTTAAACAATTTTACAATCTCTAGACAGTAGACAGTTTAT	1279		
Db	538	GCCGGAAGAGTTTATATGGATTAGCTTAAACAATTTTACAATCTCTAGACAGTAGACAGTTTAT	479		
QY	1280	AAATACTACAAATTACTTCTTTGTGAGTTATGACAGCAGTGTAAACATGTGATAAGACATG	1339		
Db	478	AAATACTACAAATTACTTCTTTGTGAGTTATGACAGCAGTGTAAACATGTGATAAGACATG	419		
QY	1340	TAAAGCCCTTTTCAGATTTCGTGCAATTTATGAATCTTGTGAATAATTTCCCTATGCAGATTGCCT	1399		
Db	418	TAAAGCCCTTTTCAGATTTCGTGCAATTTATGAATCTTGTGAATAATTTCCCTATGCAGATTGCCT	359		

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11464 row: m column: 01
High quality sequence stop: 650.
Location/Qualifiers
1. .650
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5186088"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."
BASE COUNT 205 a 97 c 116 g 232 t
ORIGIN
Query Match 35.8%; Score 631.6; DB 13; Length 650;
Best Local Similarity 99.2%; Pred. No. 5.9e-98;
Matches 645; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1046 AGAAAAACAGACCAACAATCCTGGTATCAGACATGTTTCAGTATGATCCTCGTATTATAA 1105
DB 1 AGAAAAACAGACCAACAATCCTGGTATCAGACATGTTTCAGTATGATCCTCGTATTATAA 60
QY 1106 ATTATTGGATATGTTGCGAGTATTACTTGAATCTGACAGAGCGCAATCTAAGGGAGATC 1165
DB 61 ATTATTGGATATGTTGCGAGTATTACTTGAATCTGACAGAGCGCAATCTAAGGGAGATC 120
QY 1166 CATCTCGAGCTGGAGTATATCTGACCCAGACCTACGACATTCAGATTCGAGCCGGA 1225
DB 121 CATCTCGAGCTGGAGTATATCTGACCCAGACCTACGACATTCGAGATTCGAGCCGGA 180
QY 1226 AAGTTTATATGGATTAGCTAAACAATTTACAATCTCTAGACAGTAAGCAGTTTATAAATA 1285
DB 181 AAGTTTATATGGATTAGCTAAACAATTTACAATCTCTAGACAGTAAGCAGTTTATAAATA 240
QY 1286 CTACAATTAATCTTTGTGAGTTATGACAGCAGTGAACATGTGATAAGACATGTGAAGGC 1345
DB 241 CTACAATTAATCTTTGTGAGTTATGACAGCAGTGAACATGTGATAAGACATGTGAAGGC 300
QY 1346 CTTTCAGATTTGCGCAATTAATGATCTTGATATATTTCTATGCGAGATTCCTCAACA 1405
DB 301 CTTTCAGATTTGCGCAATTAATGATCTTGATATATTTCTATGCGAGATTCCTCAACA 360
QY 1406 GCTTTTATATAAGCACAAATTAATTCAGATTTTTCAGATTTTTCGTAATAGAAAATGCTGATT 1465
DB 361 GCTTTTATATAAGCACAAATTAATTCAGATTTTTCAGATTTTTCGTAATAGAAAATGCTGATT 420
QY 1466 CTGATTCAGATCAATTTGTTGGGAATTTTACATAAATCTTTGTTAATTAATCTAGTGGGC 1525
DB 421 CTGATTCAGATCAATTTGTTGGGAATTTTACATAAATCTTTGTTAATTAATCTAGTGGGC 480
QY 1526 AAGTAGACTTCCTGCT 1585
DB 481 AAGTAGACTTCCTGCT 540

QY 1586 TCTTTATCATCTCGAATTGTATTATATATTTAAAGTGCTCATTAATAGATGATGG-ATG 1644
DB 541 TCTTTATCATCTCGAATTGTATTATATATTTAAAGTGCTCATTAATAGATGATGGCATG 600
QY 1645 TAAATGGATGTAATAATTCAGTTTATATATATATATATATCTCTAATTTGTACCC 1694
DB 601 TAAATGGATGTAATAATTCAGTTTATATATATATATATATCTCTAATTTGTACCC 650
RESULT 13
BG432589 663 bp mRNA linear EST 14-MAR-2001
LOCUS 602500747F1 NIH_MGC_75 Homo sapiens cdna clone IMAGE:4614346 5',
DEFINITION mRNA sequence.
ACCESSION BG432589
VERSION BG432589.1 GI:13339095
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 663)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1385 row: n column: 11
High quality sequence stop: 661.
Location/Qualifiers
1. .663
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4614346"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccattggcc); Site 2: SfiI (ggccattggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCCGATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
BASE COUNT 228 a 122 c 136 g 177 t
ORIGIN
Query Match 35.0%; Score 617.2; DB 12; Length 663;
Best Local Similarity 99.1%; Pred. No. 1.6e-95;
Matches 652; Conservative 0; Mismatches 3; Indels 3; Gaps 3;
QY 576 GCAGTAGCAAACTCTGGAAACCATGGCTAGATGAAGAGCTATTAGTACTTTTAAGGAAA 635
DB 1 GCAGTAGCAAACTCTGGAAACCATGGCTAGATGAAGAGCTATTAGTACTTTTAAGGAAA 60
QY 636 GGTGTTTTTATTCACAGAAAGTTACAACCTAATCCAACTTTAGGATCATGCTAAAC 695
DB 61 GGTGTTTTTATTCACAGAAAGTTACAACCTAATCCAACTTTAGGATCATGCTAAAC 120
QY 696 ACAAACTTGTACTACGGCCCAATATATGACACTGACAACTGACCCAGCCACCCAG 755
DB 121 ACAAACTTGTACTACGGCCCAATATATGACACTGACAACTGACCCAGCCACCCAG 180

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 28, 2003, 02:03:25 ; Search time 3671.46 Seconds
(without alignments)
17018.784 Million cell updates/sec
Title: US-09-981-353-104
Perfect score: 2147
Sequence: 1 ggttaggagaaagttagg.....gttactatatgttggaag 2147

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.ste.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.ste.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1992	92.8	2009	9	BC002954	BC002954 Homo sapi
2	1676.6	78.1	1823	6	AX330276	AX330276 Sequence
3	1676.6	78.1	1823	6	AX410692	AX410692 Sequence
4	1676.6	78.1	1823	9	HSU27460	U27460 Human uridi
5	1676.6	78.1	1823	11	G26908	G26908 human STS S
6	1562.2	72.8	2056	4	SSDPPGPP	X99312 S. scrofa nr
7	1491.6	69.5	2069	10	BC025585	BC025585 Mus muscu
8	1452.2	67.6	1882	10	BC026626	BC026626 Mus muscu
9	1416.2	66.0	2026	10	AF424698	AF424698 Mus muscu
10	1415	65.9	1880	10	AF004368	AF004368 Cricetulu
11	1372.8	63.9	1689	4	BOVUDPGPYR	L14019 Bovine UDP-
12	538.8	25.1	2029	3	AY069422	AY069422 Drosophill
13	523	24.4	140401	9	AC092588	AC092588 Homo sapi
14	504.4	23.5	607	9	HUMZA06D10	AF086113 Homo sapi
15	448.4	20.9	1625	3	AF150929	AF150929 Dictyoste
16	419.8	19.6	671	6	AX401201	AX401201 Sequence
17	413.8	19.3	1419	6	AX489345	AX489345 Sequence
18	381.6	17.8	1722	8	AF203909	AF203909 Musa acum
19	376.6	17.5	2488	8	AF100788	AF100788 Gracilari
20	376.2	17.5	2711	8	SCYKL037W	Z28035 S. cerevisia
21	375.2	17.5	12399	8	SCX112	X69584 S. cerevisi
22	375.2	17.5	396	6	AX093225	AX093225 Sequence
23	374	17.4	1398	6	E10417	E10417 UDP Glucose
24	369.4	17.2	1831	8	AF281081	AF281081 Astragalu
25	369.4	17.2	1831	8	AF435969	AF435969 Amorpha f
26	366	17.0	42070	8	SPCC1322	AL035259 S.pombe c
27	358.4	16.7	1783	8	HVUDPGPP	X91347 H. vulgare m
28	357.8	16.7	1413	8	AY040042	AY040042 Arabidops
29	357.8	16.7	1628	8	AF360244	AF360244 Arabidops
30	357.8	16.7	1800	8	AF361816	AF361816 Arabidops
31	356.6	16.6	1699	8	AB013353	AB013353 Pyrus pyr
32	354.8	16.5	1746	8	AF249880	AF249880 Oryza sat
33	352.8	16.4	1441	8	AY059148	AY059148 Arabidops
34	352.8	16.4	1795	8	AY035071	AY035071 Arabidops
35	350.8	16.3	1683	8	STUDPGVP	Z18924 S. tuberosum
36	350.8	16.3	1758	8	POTUDPGP	D0667 Potato mRNa
37	349.6	16.3	1739	8	AF361605	AF361605 Arabidops
38	346	16.1	1807	8	AB069659	AB069659 Pyrus pyr
39	344.4	16.0	1742	8	AB069658	AB069658 Pyrus pyr
40	339.8	15.8	1700	8	AB062606	AB062606 Oryza sat
41	333.6	15.5	39565	3	CEK08E3	Z81568 Caenorhabdi
42	325.4	15.2	1128	8	AB055502	AB055502 Nicotiana
43	319.2	14.9	1954	8	AB037186	AB037186 Candida g
44	301	14.0	127561	9	AC114748	AC114748 Homo sapi
45	287.4	13.4	194439	3	CNS07EGA	AL590443 chromosom

ALIGNMENTS

RESULT 1
BC002954
LOCUS
DEFINITION Homo sapiens, UDP-glucose pyrophosphorylase 2, clone MSC:1251
ACCESSION BC002954
VERSION BC002954.1 GI:12804192
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2009)
AUTHORS Strausberg,R.
TITLE Direct Submission

JOURNAL

Submitted (06-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutache, Oliver Lee, Soo Nee Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Farvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasya van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 8 Row: e Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13027637.

FEATURES

source

1..2009
/organism="Homo sapiens"
/db_xref="LocustID:7360"
/db_xref="taxon:9606"
/clone="MGC:1251 IMAGE:3542849"
/issue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH MGC 8"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
164..1657
/codon_start=1
/product="UDP-glucose pyrophosphorylase 2"
/protein_id="AAH02954.1"
/db_xref="GI:12804193"
/translation="MSQDASQFQSVIRQELSLVKKLEKLTITASHEPHTKKDL
DGRFLKRFLEKPSQVDMGKIQRPPDSQPEYKIKARGLPDMISSVLNKLVVVKL
NGGLTSGCKGPKSLIGVRNENTFLDLTVQOIEHLNKTNDYDPLVNLNFTDDET
KKILQYHCRVYITFQNSRYPRIKESLLPVAKDVSYSSENTAEPYPPGHGDIYAS
FVNSGLLDTFIEGKEYIFVSNIDNLGATVDLYILNLMNPNPKRCFVMEVNTKTR
ADVKGTLTOYEGKRLVEIAOVKRAHVDKFSKFKFINNNLWISLAAVKRLQEQ
NAIDWEIIVNAKTLDGLNVIQLETAAGVAIKSFNSLGINVPSRFLPVKTTSLLL
VNSNLSLNAGSLTMSKREPTFVLKLGSSFTKQDYLRPFESIPDMLDLHTVS
GDVTFGNVSLGTGTVIIITANHGRIDIDPPGAVLENKIVSGNRLRLDH"

CDS

666 a 360 c 439 g 544 t
Query Match 92.8%; Score 1992; DB 9; Length 2009;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT
ORIGIN

78 GCGACCGTTAGGAGGGGCTCAAGGTGTGCATGTGTGAGGGAAGAGAGAGAGAGAGAGG 137
1 GCGACCGTTAGGAGGGGCTCAAGGTGTGCATGTGTGAGGGAAGAGAGAGAGAGAGAGG 60
138 GCGCCTCAGAGTGACTTTCAGCCTCGAGCCTTCTCCGGGGGCGCCATAACGCCCC 197
61 GCGCCTCAGAGTGACTTTCAGCCTCGAGCCTTCTCCGGGGGCGCCATAACGCCCC 120
198 AATTTCCAGCTGCTAAAGGAAGAGAGAGATCTTAGCAAGCAATGTCTCAAGATGGTGC 257
121 AATTTCCAGCTGCTAAAGGAAGAGAGAGATCTTAGCAAGCAATGTCTCAAGATGGTGC 180
258 TTCTCAGTTTCCAGAGTCAATTCGGCAAGAGCTAGAAATTTATCTGTGAAGAAGCACTAGA 317
181 TTCTCAGTTTCCAGAGTCAATTCGGCAAGAGCTAGAAATTTATCTGTGAAGAAGCACTAGA 240

318 AAAAATCTCACCACAGCATCATCATGAATTTTGGACACACCAAAAAAGACCTGGATGG 377
241 AAAAATCTCACCACAGCATCATCATGAATTTTGGACACACCAAAAAAGACCTGGATGG 300
378 ATTTCCGGAAGCTATTTTATAGATTTTTCACAGAAAAGGGGCTTCTGTGGATTTGGGAAA 437
301 ATTTCCGGAAGCTATTTTATAGATTTTTCACAGAAAAGGGGCTTCTGTGGATTTGGGAAA 360
438 AATCAGAGAGCCCTCGAGATTCGATTCACCCCTATGAAAAGATTAAGGCGCAGGGGCTT 497
361 AATCAGAGAGCCCTCGAGATTCGATTCACCCCTATGAAAAGATTAAGGCGCAGGGGCTT 420
498 GCCTGATAATATATCTTCCGTTGTCACAACTAGTGTGTGAACTCAATTTGGTGGTTT 557
421 GCCTGATAATATATCTTCCGTTGTCACAACTAGTGTGTGAACTCAATTTGGTGGTTT 480
558 GGGAAACAGCATGGGTGTCACAAAGCCCTTAAAGTCTGATTTGGTGTGAGGAATGAGAAATAC 617
481 GGGAAACAGCATGGGTGTCACAAAGCCCTTAAAGTCTGATTTGGTGTGAGGAATGAGAAATAC 540
618 CTTCCTGATCTGACTGTTCAGCAAAATTTGAACATTTGAATTAACCTACAATACAGATGT 677
541 CTTCCTGATCTGACTGTTCAGCAAAATTTGAACATTTGAATTAACCTACAATACAGATGT 600
678 TCCTCTTGTGTTTAAATGAATCTTTTAAACACGAGTGAAGATACCAAAAAAATCTACAGAA 737
601 TCCTCTTGTGTTTAAATGAATCTTTTAAACACGAGTGAAGATACCAAAAAAATCTACAGAA 660
738 GTACAAATCATTTGCTGTGAAATCTACATTTCAATCAAAAGAGTACCCGAGGATTA 797
661 GTACAAATCATTTGCTGTGAAATCTACATTTCAATCAAAAGAGTACCCGAGGATTA 720
798 TAAAGAAATCTTTACTTCTCTAGCAAAAGAGCTGTCTTACTCAGGGGAAAAATACAGAGC 857
721 TAAAGAAATCTTTACTTCTCTAGCAAAAGAGCTGTCTTACTCAGGGGAAAAATACAGAGC 780
858 TTGGTACCCCTCAGGTCAATGTTGATTTTACGCGCATTTTCTACAACTCTGATTCCTTGA 917
781 TTGGTACCCCTCAGGTCAATGTTGATTTTACGCGCATTTTCTACAACTCTGATTCCTTGA 840
918 TACCTTTATAGAGAGGCAAAAGAGTATATTTTGTGTCTAAACATAGATAATCTCGGTGC 977
841 TACCTTTATAGAGAGGCAAAAGAGTATATTTTGTGTCTAAACATAGATAATCTCGGTGC 900
978 CACAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCACCAATGAAAACCGCTGTGA 1037
901 CACAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCACCAATGAAAACCGCTGTGA 960
1038 ATTTGTCAATGAGTCAAAATTAACACAGTGCAGATGTAAGGCGGAGCACTCACTCA 1097
961 ATTTGTCAATGAGTCAAAATTAACACAGTGCAGATGTAAGGCGGAGCACTCACTCA 1020
1098 ATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCACAAAGCACATGTAGACGA 1157
1021 ATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCACAAAGCACATGTAGACGA 1080
1158 GTTCAAGTCTGTATCAAAAGTTCAAAATATTTAATAACAAACACCTATGATTTCTCTTGC 1217
1081 GTTCAAGTCTGTATCAAAAGTTCAAAATATTTAATAACAAACACCTATGATTTCTCTTGC 1140
1218 AGCAGTAAAAAGACTGACAGGAGCAAAATGCAATTCACATGCAATTCATTTGTAATGCAAA 1277
1141 AGCAGTAAAAAGACTGACAGGAGCAAAATGCAATTCACATGCAATTCATTTGTAATGCAAA 1200
1278 GACTTTGGATGGAGCCTTGAATGTCAATTCATTTAGAAAATGCAAGTGGGGCTGCCATCAA 1337
1201 GACTTTGGATGGAGCCTTGAATGTCAATTCATTTAGAAAATGCAAGTGGGGCTGCCATCAA 1260
1338 AAGTTTTCAGAAATCTCTAGGTATTAATGTCGAGGAGGCGGTTTCTTGCCTGTCAAAC 1397
1261 AAGTTTTCAGAAATCTCTAGGTATTAATGTCGAGGAGGCGGTTTCTTGCCTGTCAAAC 1320

Db 1113 TGCATTTGACATGAAATCATTTGTAATGCAAGAGCTTTGGATGAGCGCTGAATGTGAT 1172
 Qy 1305 TCAATTTAGAACTGCAAGTGGGCTGCCATCAAAAGTTTGAAGATTTCTAGGTATTA 1364
 Db 1173 TCAATTTAGAACTGCAAGTGGGCTGCCATCAAAAGTTTGAAGATTTCTAGGTATTA 1232
 Qy 1365 TGTGCCAAGGAGCGGTTTCTGCTGTCAAAACACATCAGATCTTTGCTGGTGATGTC 1424
 Db 1233 TGTGCCAAGGAGCGGTTTCTGCTGTCAAAACACATCAGATCTTTGCTGGTGATGTC 1292
 Qy 1425 AAACCTCTATAGCTTAATGAGGAGTCTGCAATGAGTGAAAGCGGGAATTTCTTAC 1484
 Db 1293 AAACCTCTATAGCTTAATGAGGAGTCTGCAATGAGTGAAAGCGGGAATTTCTTAC 1352
 Qy 1485 AGTGCCTTGGTTAAATAGGAGTCTTTTACGAAGTTTCAAGATTTCTAAGAAGATT 1544
 Db 1353 AGTGCCTTGGTTAAATAGGAGTCTTTTACGAAGTTTCAAGATTTCTAAGAAGATT 1412
 Qy 1545 TGAAGATATACAGATATGCTTGAATGGATCACTCACATTTTCAAGAGATGTGACATT 1604
 Db 1413 TGAAGATATACAGATATGCTTGAATGGATCACTCACATTTTCAAGAGATGTGACATT 1472
 Qy 1605 TGAAGATATGCTTCAATTAAGGAGCAAGGTTATCATCATTTGCAATCATGTCACAGAT 1664
 Db 1473 TGAAGATATGCTTCAATTAAGGAGCAAGGTTATCATCATTTGCAATCATGTCACAGAT 1532
 Qy 1665 TGATATCCACCTGGAGCAGTATTAGAGAACAGATTTGCTCGAAACCTTCGCAATCTT 1724
 Db 1533 TGATATCCACCTGGAGCAGTATTAGAGAACAGATTTGCTCGAAACCTTCGCAATCTT 1592
 Qy 1725 GGACCACTGAAATGAAATTAAGTGTGACACTTAAATTAATGGCTAGTTTCTTACAAATGA 1784
 Db 1593 GGACCACTGAAATGAAATTAAGTGTGACACTTAAATTAATGGCTAGTTTCTTACAAATGA 1652
 Qy 1785 AATGTTCTTAGGATTTCAATTAAGGAGTCTTACTATGTTTACTGACCTCGAGTG 1844
 Db 1653 AATGTTCTTAGGATTTAGGACACTTAAAGGATCTTTACTATGTTTACTGACCTCGAGTG 1712
 Qy 1845 TTGATTTTAAATAGAGTTTCTGCAAGTATGCTTTTACTGTAAGAAAGACAGATGGA 1904
 Db 1713 TTGATTTTAAATAGAGTTTCTGCAAGTATGCTTTTACTGTAAGAAAGACAGATGTT 1772
 Qy 1905 GCAATCTTCTTCTTTGAAGAGATCCCAAGTGTCTTCTTCTTAA 1953
 Db 1773 GCAATCTTCTTCTTTGAAGAGATCCCAAGTGTCTTCTTCTTAA 1819

RESULT 4
 HSU27460
 LOCUS
 DEFINITION Human uridine diphosphoglucose pyrophosphorylase mRNA, complete cds.
 ACCESSION U27460.1 GI:881393
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1823)
 Dugleby, R.G., Chao, Y.C., Huang, J.G., Peng, H.I., and Chang, H.Y.
 Sequence differences between human muscle and liver cDNAs for
 UDPglucose pyrophosphorylase and kinetic properties of the
 recombinant enzymes expressed in *Escherichia coli*
 Eur. J. Biochem. 235 (1-2), 173-179 (1996)
 96202932
 8631325
 2 (bases 1 to 1823)
 Chang, H.-Y.
 Direct Submission
 Submitted (23-MAY-1995) Hwan-You Chang, Molecular and Cellular
 Biology, Chang-Gung College of Medicine and Technology, 259 Wen-Hwa

FEATURES
 source
 First Road, Kwei-San, Tao-Yuan, Taiwan
 Location/Qualifiers
 1..1823
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="pHC377"
 /sex="female"
 /tissue type="skeletal muscle"
 /clone_lib="Stratagene # 936215"
 1..108
 5'UTR
 CDS
 109..1602
 /EC_number="2.7.7.9"
 /function="synthesis of UDP-glucose from
 glucose-1-phosphate"
 /note="similar to uridine diphosphoglucose
 pyrophosphorylase in human liver, Swiss-Prot Accession
 Number Q07131; the 5'UTR and 3'UTR of this clone are
 completely different from those of the liver form"
 /codon_start=1
 /product="uridine diphosphoglucose pyrophosphorylase"
 /protein_id="AAB05640.1"
 /db_xref="GI:881394"
 /translation="MSQDQASQFQEVIRQELSLVKELEKILTTASSHEFEHTKKDL
 DGPRLFRFLQKPSVDMGKIQRPPDSIQYEKI KARGLPDNISSVLKLVVVKL
 NGGLTSMGCKGPKSLFNQSRPRINKESLRPAKDVSYSGENTEARWYPGHGDIYAS
 FVNSGLDITFEGEKEYIFVSNIDNLGATVDLVILANHLINPPNKRCEFFVMEVNTKTR
 ADVKGGTITQVEGKLRLVEIAOVPAKHVDPEKSVKFKFNNTNNLWISLAARLRQEQ
 NADMEIIVNAKTLDCGLAVIOLETAVGAIAIKFNSLGINVPRSLFVKTSLDL
 VMSNLXSLNAGSLTMEKREFPVPLVKGSSFTKVDYLRREFSIPDMLDLHLTVS
 GDVTFGKNSLKGTVIIIIANHNDRIIDIPPGAVLENKIVSGNLRILDH"
 1603..1823
 3'UTR
 BASE COUNT 593 a 337 c 402 g 491 t
 ORIGIN
 Query Match 78.1%; Score 1676.6; DB 9; Length 1823;
 Best Local Similarity 98.5%; Pred.No. 0;
 Matches 1703; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
 Qy 225 AGATCTTAGCAAGCAATGTCTCAAGATGGTGTCTCTCAGTTCCAAAGAGTCAATCGGCA 284
 Db 93 AGATCTTAGCAAGCAATGTCTCAAGATGGTGTCTCTCAGTTCCAAAGAGTCAATCGGCA 152
 Qy 285 AGAGCTAGAGTATCTGTGAAGAGCACTAGAAAAATACTCACACAGCATCATCA 344
 Db 153 AGAGCTAGAGTATCTGTGAAGAGCACTAGAAAAATACTCACACAGCATCATCA 212
 Qy 345 TGAATTTGAGCACACCAAAAAAGACCTGGATGGATTTTCGGAAGCTATTTCTAGATTTTT 404
 Db 213 TGAATTTGAGCACACCAAAAAAGACCTGGATGGATTTTCGGAAGCTATTTCTAGATTTTT 272
 Qy 405 GCAAGAAAAGGGGCTTCTGTGGATTTGGGAAAAATCCAGAGACCCCTGAAGATTCAT 464
 Db 273 GCAAGAAAAGGGGCTTCTGTGGATTTGGGAAAAATCCAGAGACCCCTGAAGATTCAT 332
 Qy 465 TCAACCTATGAAAAGATAAAGGCCAGGGCTTCGCTGATATATATATATATATATATATAT 524
 Db 333 TCAACCTATGAAAAGATAAAGGCCAGGGCTTCGCTGATATATATATATATATATATATAT 392
 Qy 525 CAAACTAGTGGTGAACCTCAATGGTGGTGGGAAACCAAGCATGGGCTGCAAGGCC 584
 Db 393 CAAACTAGTGGTGAACCTCAATGGTGGTGGGAAACCAAGCATGGGCTGCAAGGCC 452
 Qy 585 TAAAGTCTGATTTGGTGTGAGGAATGAGATACCTTTCTGGATCTGATCTGATCTGATCTGAT 644
 Db 453 TAAAGTCTGATTTGGTGTGAGGAATGAGATACCTTTCTGGATCTGATCTGATCTGATCTGAT 512
 Qy 645 TGAACATTTGAATAAAACCTACATACAGATGTTCTCTTTGTTTAAATGAATCTTTTAA 704
 Db 513 TGAACATTTGAATAAAACCTACATACAGATGTTCTCTTTGTTTAAATGAATCTTTTAA 572
 Qy 705 CACGGATGAAGATACCAAAAAATACTACAGAGTCAATCATTTGCTGCTGGAATCTA 764

Db 573 CAGGATGAAGATACCAAAAAAATCTACAGAAGTACAATCATCTCTGCTGAAAAATCTA 632
Qy 765 CACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAACTTTACTTCTCTGTAGCAAA 824
Db 633 CACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAACTTTACTTCTGTAGCAAA 692
Qy 825 GGAGCTGTCTACTCAGGGGAAAAATACAGAGCTTGGTACCTCCAGGTCTATGGTGATAT 884
Db 693 GGAGCTGTCTACTCAGGGGAAAAATACAGAGCTTGGTACCTCCAGGTCTATGGTGATAT 752
Qy 885 TTACGCCAGTTTCTCAAACTCTGGATTCTTGATACCTTTATAGAGAAGCAAGAGTA 944
Db 753 TTACGCCAGTTTCTCAAACTCTGGATTCTTGATACCTTTATAGAGAAGCAAGAGTA 812
Qy 945 TATTTTGTGTCTAATCATAGATAATCTGGGTGCCACAGTGGATCTGTATATCTTAAATCA 1004
Db 813 TATTTTGTGTCTAATCATAGATAATCTGGGTGCCACAGTGGATCTGTATATCTTAAATCA 872
Qy 1005 TCTAATGAACCCACCACCAATGGAACCGCTGTGAATTTGTCTGGAAGTCAACAATAAATAC 1064
Db 873 TCTAATCAACCCACCACCAATGGAACCGCTGTGAATTTGTCTGGAAGTCAACAATAAATAC 932
Qy 1065 ACGTGCAGATGTAAAGGGCGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGA 1124
Db 933 ACGTGCAGATGTAAAGGGCGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGA 992
Qy 1125 AATTGCTCAAGTCCAAAAGCAGATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAAT 1184
Db 993 AATTGCTCAAGTCCAAAAGCAGATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAAT 1052
Qy 1185 ATTTAATCAAAACCACTATGATTTCTTTCGACAGTTAAAGACATGACAGGAGCAAAA 1244
Db 1053 ATTTAATCAAAACCACTATGATTTCTTTCGACAGTTAAAGACATGACAGGAGCAAAA 1112
Qy 1245 TGCCATTGACATGGAATCATTTGTAATGCAAGACTTTTGGATGAGGCTGAAATGTCAT 1304
Db 1113 TGCCATTGACATGGAATCATTTGTAATGCAAGACTTTTGGATGAGGCTGAAATGTCAT 1172
Qy 1305 TCAATTAGAACTGACATGAGGCTGCCATCAAAAGTTTGGAAATTTCTAGGTATTAA 1364
Db 1173 TCAATTAGAACTGACATGAGGCTGCCATCAAAAGTTTGGAAATTTCTAGGTATTAA 1232
Qy 1365 TGTGCCAAGGAGCGGTTTTCTGCTGTCGCAACACATCAGATCTTCTGCTGGTATGTC 1424
Db 1233 TGTGCCAAGGAGCGGTTTTCTGCTGTCGCAACACATCAGATCTTCTGCTGGTATGTC 1292
Qy 1425 AAACCTCTATAGTCTTAATGCAAGATCTCTGCAATGATGAAAGCGGGAATTTCTTAC 1484
Db 1293 AAACCTCTATAGTCTTAATGCAAGATCTCTGCAATGATGAAAGCGGGAATTTCTTAC 1352
Qy 1485 AGTGCCCTTGGTTAAATTAGGCAGTTCTTTTACGAAGGTTCAAGATTATCTAAGAGATT 1544
Db 1353 AGTGCCCTTGGTTAAATTAGGCAGTTCTTTTACGAAGGTTCAAGATTATCTAAGAGATT 1412
Qy 1545 TGAAGATATACAGATATGCTTGAATTTGATCTGATCACTCAGTTTCAGGAGATGTACATT 1604
Db 1413 TGAAGATATACAGATATGCTTGAATTTGATCTGATCACTCAGTTTCAGGAGATGTACATT 1472
Qy 1605 TGGAAAAATGTTTCATTAAAGGGGAACGGTTATCATATTGCAAAATCATGTGTACAGAA 1664
Db 1473 TGGAAAAATGTTTCATTAAAGGGGAACGGTTATCATATTGCAAAATCATGTGTACAGAA 1532
Qy 1665 TGATATCCCACTGAGCAGATATTAGAGNACAAGATTGTGTCTGAAACCTTCGCATCTT 1724
Db 1533 TGATATCCCACTGAGCAGATATTAGAGNACAAGATTGTGTCTGAAACCTTCGCATCTT 1592
Qy 1725 GGACCACTGAATGAAAAATCTGTGGACACTTAAATAATGGCTAGTTTCTTACAATGA 1784
Db 1593 GGACCACTGAATGAAAAATCTGTGGACACTTAAATAATGGCTAGTTTCTTACAATGA 1652
Qy 1785 AATGTTCTTAGGATTTCTAAATATAGGCAAGGTACTTTACTATGTTACTGTACCTCAGTG 1844
Db 1653 AATGTTCTTAGGATTTAGGCACTAAAGGTACTTTACTATGTTACTGTACCTCAGTG 1712

Qy 1845 TTGATTTTAAATAGAGTTTTTCTGCAGTAGTCTTTTAGTCTTAAGAAAACACAGATGGA 1904
Db 1713 TTGATTTTAAATAGAGTTTTTCTGCAGTAGTCTTTTAGTCTTAAGAAAACACAGATGTT 1772
Qy 1905 GCAATACTTTCTTCTTTTGAAGAGAAATCCCAAAAGTTAGTTTCATCTTAA 1953
Db 1773 GCAATACTTTCTTCTTTTGAAGAGAA--TCCCAAGTTAGTTACTCTTAA 1819

RESULT 5

G26908
LOCUS human STS SHGC-31592, sequence tagged site.
DEFINITION human STS SHGC-31592, sequence tagged site.
ACCESSION G26908
VERSION G26908.1 GI:1375158
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1823)
AUTHORS Myers, R.M.
JOURNAL Unpublished (1995)
COMMENT

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@hgc.stanford.edu

Primer A: GAAATGAAAATACTGTGGACACTT
Primer B: TTTGGATCTCTTCAAGAAGG
STS size: 204
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:

MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from U27460
-- Washington University/Merck EST sequence.

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="2"
STS
primer_bind
1601..1804
primer_bind
1601..1625
complement(1783..1804)
BASE COUNT 593 a 337 c 402 g 491 t
ORIGIN

Query Match 78.1%; Score 1676.6; DB 11; Length 1823;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1703; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

OY 225 AGATCTTAGCAAGCAATGTCTCAAGATGGTCTCTCAGTTCCTCAAGAGTCAATTCGGCA 284
DB 93 AGATCTTAGCAAGCAATGTCTCAAGATGGTCTCTCAGTTCCTCAAGAGTCAATTCGGCA 152
OY 285 AGAGCTAGCAATTAATCTGTGAAGAGAGAACTAGAGAAATATCTCACCACAGCATCATCACA 344
DB 153 AGAGCTAGCAATTAATCTGTGAAGAGAGAACTAGAGAAATATCTCACCACAGCATCATCACA 212
OY 345 TGAATTTGAGCACACCAACCAAGACCTGGATGGATTTTCGAGAGCTATTTTCATAGATTTT 404
DB 213 TGAATTTGAGCACACCAACCAAGACCTGGATGGATTTTCGAGAGCTATTTTCATAGATTTT 272
OY 405 GCAAGAAAAGGGGCTTCTGTGGATTTGGGAAAATCCAGAGACCCCTCGAAGATTCGAT 464
DB 273 GCAAGAAAAGGGGCTTCTGTGGATTTGGGAAAATCCAGAGACCCCTCGAAGATTCGAT 332
OY 465 TCAACCTTATGAAAGATTAAGCCAGGGGCTTTCCTGATTAATATATATCTTCCTGTGAA 524
DB 333 TCAACCTTATGAAAGATTAAGCCAGGGGCTTTCCTGATTAATATATATCTTCCTGTGAA 392
OY 525 CAAACTAGTGGTGGTGAACCTCAATGGTGGTGGGAAACCCAGCATGGCTGCAAGGCC 584
DB 393 CAAACTAGTGGTGGTGAACCTCAATGGTGGTGGGAAACCCAGCATGGCTGCAAGGCC 452
OY 585 TAAAGTCTGATTTGGTGGTGAAGATGAGAAATACCTTTCTGGATCTGATCTGTTGAGCAAT 644
DB 453 TAAAGTCTGATTTGGTGGTGAAGATGAGAAATACCTTTCTGGATCTGATCTGTTGAGCAAT 512
OY 645 TGAACATTTGAATAAAGACCTTACAAATACAGATGTTCTCTGTTTAAAGAACTCTTTAA 704
DB 513 TGAACATTTGAACAAAGACCTTACAAATACAGATGTTCTCTGTTTAAAGAACTCTTTAA 572
OY 705 CACGGATGAAGATACCAAAAAATACACAGAGTACAATCATTTGCTGATCTGATCTGTTGAGCAAT 764
DB 573 CACGGATGAAGATACCAAAAAATACACAGAGTACAATCATTTGCTGATCTGATCTGTTGAGCAAT 632
OY 765 CACTTTCAATCAAGAGTACCCGAGGATTAATAAGAACTCTTACTTCTGTTAGCAAA 824
DB 633 CACTTTCAATCAAGAGTACCCGAGGATTAATAAGAACTCTTACTTCTGTTAGCAAA 692
OY 825 GGAAGTCTTACTCAGGGGAAAATACAGAACTTGGTACCTCCAGTCTGATCTGATCTGAT 884
DB 693 GGAAGTCTTACTCAGGGGAAAATACAGAACTTGGTACCTCCAGTCTGATCTGATCTGAT 752
OY 885 TTACGCCAGTTTCTACAACTCTGGATTTGCTGATACCTTTATAGAGAGCAAGAGTA 944
DB 753 TTACGCCAGTTTCTACAACTCTGGATTTGCTGATACCTTTATAGAGAGCAAGAGTA 812
OY 945 TATTTTGTCTTAACATAGATAATCTGGTGGCAAGTGGATCTGATCTGATCTGATCTGAT 1004
DB 813 TATTTTGTCTTAACATAGATAATCTGGTGGCAAGTGGATCTGATCTGATCTGATCTGAT 872
OY 1005 TCTAATGAACCCCAATGGAACACGCTGTGAATTTGCTGATGGAAGTCAAAATAAAT 1064
DB 873 TCTAATGAACCCCAATGGAACACGCTGTGAATTTGCTGATGGAAGTCAAAATAAAT 932
OY 1065 AGTGCAGATGAAGGGGGGACACTCACTAATATGAAGGCAAACTGAGCTGGTGA 1124
DB 933 AGTGCAGATGAAGGGGGGACACTCACTAATATGAAGGCAAACTGAGCTGGTGA 992
OY 1125 AATTGCTCAAGTGCCAAAAGCACATGTAGACGAGTCAAGTCTGATCAAGTTCAAAAT 1184
DB 993 AATTGCTCAAGTGCCAAAAGCACATGTAGACGAGTCAAGTCTGATCAAGTTCAAAAT 1052
OY 1185 ATTTAATCAAAACCACTATGGATTTCTTTCGACGAGTTAAAGAGCTGACGAGCAAA 1244
DB 1053 ATTTAATCAAAACCACTATGGATTTCTTTCGACGAGTTAAAGAGCTGACGAGCAAA 1112
OY 1245 TGCCATTGACATGAAATCATTTGGAATGCAAGACTTTGGATGGAGGCTGATGTCAT 1304
DB 1113 TGCCATTGACATGAAATCATTTGGAATGCAAGACTTTGGATGGAGGCTGATGTCAT 1172

OY 1305 TCAATTAGAAAACGAGTAGGGCTGCCATCAAAAGTTTTCGAGATTTCTTAGGTATTAA 1364
DB 1173 TCAATTAGAAAACGAGTAGGGCTGCCATCAAAAGTTTTCGAGATTTCTTAGGTATTAA 1232
OY 1365 TGTGCAAGAGAGCCGTTTCTGCTGTCAAAACCAATCAGATCTCTTCTGCTGGTATGTC 1424
DB 1233 TGTGCAAGAGAGCCGTTTCTGCTGTCAAAACCAATCAGATCTCTTCTGCTGGTATGTC 1292
OY 1425 AAACCTCTATAGTCTTAATGAGGATCTCTCAGATGAGTCAAAAGCGGGAATTTCTCTAC 1484
DB 1293 AAACCTCTATAGTCTTAATGAGGATCTCTCAGATGAGTCAAAAGCGGGAATTTCTCTAC 1352
OY 1485 AGTGCCCTTGGTTAAATTTAGGAGTCTTTTACGAAGGTTCAAGATTTCTTAAGAAGATT 1544
DB 1353 AGTGCCCTTGGTTAAATTTAGGAGTCTTTTACGAAGGTTCAAGATTTCTTAAGAAGATT 1412
OY 1545 TGAAGATATACAGATATGTTGAATTTGGATCACTCAGATTTTCAGAGATGTGACATT 1604
DB 1413 TGAAGATATACAGATATGTTGAATTTGGATCACTCAGATTTTCAGAGATGTGACATT 1472
OY 1605 TGGAAAAATGTTTCAATAAGGGAACGGTTATCATCATTTGCAAAATCATGGTACAGAA 1664
DB 1473 TGGAAAAATGTTTCAATAAGGGAACGGTTATCATCATTTGCAAAATCATGGTACAGAA 1532
OY 1665 TGATATCCACCTGGAGCAGTATTAGAGAACAGATTTGCTCTGGAACCTTCGCATCTT 1724
DB 1533 TGATATCCACCTGGAGCAGTATTAGAGAACAGATTTGCTCTGGAACCTTCGCATCTT 1592
OY 1725 GGACCACTGAAATGAAAAATCTGTGGACACTTTAAATAATGGCTAGTTTCTTACAATGA 1784
DB 1593 GGACCACTGAAATGAAAAATCTGTGGACACTTTAAATAATGGCTAGTTTCTTACAATGA 1652
OY 1785 AATGTTCTTAGGATTTAAATAGGAGTACTTTTACTATGTTACTGACCTGCAGTG 1844
DB 1653 AATGTTCTTAGGATTTAGGACACTAAAGGTTACTTTTACTATGTTACTGACCTGCAGTG 1712
OY 1845 TTGATTTTAAATAGAGTTTCTGAGATGCTTTTACTGATGTTACTGATGTTACTGATGTT 1904
DB 1713 TTGATTTTAAATAGAGTTTCTGAGATGCTTTTCTGATGTTACTGATGTTACTGATGTT 1772
OY 1905 GCAATACTTTCTTCTTTGAAGAGAAATCCAAAAGTTAGTTTCAATCTTAA 1953
DB 1773 GCAATACTTTCTTCTTTGAAGAGAAATCCAAAAGTTAGTTTCAATCTTAA 1819

RESULT 6

SSUPDGP

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

SSUPDGP 2056 bp mRNA linear MAM 21-DEC-1996
S.scrofa mRNA for UDP glucose pyrophosphorylase.

X99312.1 GI:1752676

UDP-glucose pyrophosphorylase.

Sus scrofa.

Sus scrofa.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

Loof, C. and Paul, S.

CDNA sequencing of the porcine UDP glucose pyrophosphorylase gene

Unpublished

Loof, C.

Direct Submission

Submitted (24-JUN-1996) C. Loof, Institute of Animal Breeding,

Agriculture, University of Kiel, Olshausenstrasse 40, 24118 Kiel,

FRG

Location/Qualifiers

1. 2056

/organism="Sus scrofa"

/db_xref="taxon:9823"

/tissue_type="muscle"

/dev_stage="adult"

181..1707

1150 GTAGACGAGTCAAGTCTGTATCAAGTTCAAATATTTAATACAAACCAACCTATGGATT 1209
 1122 GTTGAATAGTTCAGTCTGTGTCAAAATTTAAATATTTAATACAAACCAACCTATGGATT 1181
 1210 TCTCTTGAGCAGGATTTAAAGAGCTGCAGGAGCAAAATGCCATTGACATGGAATCATTTGTG 1269
 1182 TCTCTTGAGCAGGATTTAAAGAGCTGCAGGAGCAAAATGCCATTGACATGGAATCATTTGTG 1241
 1270 AATGCAAGACATTTGGATGGAGGCTGAATGTCATCAATTTAGAACTGCGATAGGCGCT 1329
 1242 AATCAAGACATTTGGATGGAGGCTGAATGTCATCAATTTAGAACTGCGATAGGCGCT 1301
 1330 GCCATCAAGATTTGAGAAATCTCTAGTATTAATGTCGCAAGGAGCGCTTTCTGCT 1389
 1302 GCAATTAAGATTTGAGAAATCTCTAGTATTAATGTCGCAAGGAGCGCTTTCTGCT 1361
 1390 GTCAAAACCAATCAGATCTCTGCTGCTGATGTCGCAAGGAGCGCTTTCTGCT 1449
 1362 GTGAAACCAATCAGATCTCTGCTGCTGATGTCGCAAGGAGCGCTTTCTGCT 1421
 1450 TCTCTGACATGAGTGAAGGCGGAAATTTCTCAGTGCCTTGGTTAAATTTAGGCGCT 1509
 1422 TCTCTGACATGAGTGAAGGCGGAAATTTCTCAGTGCCTTGGTTAAATTTAGGCGCT 1481
 1510 TCTTTTACGAGGTTCAAGATTTCTAGAGATTTGAAAGATTTGAAAGATTTGAAAGATTTGAA 1569
 1482 TCTTTTACGAGGTTCAAGATTTCTAGAGATTTGAAAGATTTGAAAGATTTGAAAGATTTGAA 1541
 1570 TTGGATCACCTCACAGTTTTCAGGAGATGTGACATTTGGAAATTTGTTTCAATTAAGGGA 1629
 1542 CTGGACCACTCACTGTTTCAGGAGATGTGACATTTGGAAATTTGTTTCAATTAAGGGA 1601
 1630 ACGTTTATCATGTTGCAATCATGTTGAGAGATTTGAAAGATTTGAAAGATTTGAAAGATTTGAA 1689
 1602 ACAGTTTATTTATTTGCAATCATGTTGAGAGATTTGAAAGATTTGAAAGATTTGAAAGATTTGAA 1661
 1690 GAGAACAGATTTGTTGTTGCAAGCTTTGCAAGCTTTGCAAGCTTTGCAAGCTTTGCAAGCTTTGCA 1749
 1662 GAGAACAGATTTGTTGTTGCAAGCTTTGCAAGCTTTGCAAGCTTTGCAAGCTTTGCAAGCTTTGCA 1721
 1750 GGACACTT-----AAATAATGGGCT-----AGTTTCTTCAATTAAGATTTGTTTCTAGGATT 1800
 1722 ATACACTTCTACTAATTTAGGCTTAAAGAGTTTCTTCTAATAAATAATATCTCTAAGATT 1781
 1801 CTAAATAGCAGGATTTACTTTA-CTATGTTACTGTACCTTGCAGTGTGATTTTAAATA 1859
 1782 ATAAATAAGCAGGATTTACTTTTACTATGTTGTTACTCTGCAAGTGTGATTTTAAATA-G 1840
 1860 GAGTTTCTGCGATGCTTTTGTCTAAGAAAGCAGATGGAGCAATCTTTCTCTTC 1919
 1841 TAGTTTCTGCAATGAGCTTTTAGTCTAAGGAAAGCATAGATGAAGCAATCTTTCTCTTC 1900
 1920 TTTGAGAGAAATCCCAAAAGTTAGTTTCTTAAAGTGCAATTTGTTTAAATCTTAAAC 1979
 1901 TTTGAGAAAGCTTAAATTTAGTTTCTTAAAGTGCAATTTGTTTAAATCTTAAAC 1960
 1980 TGGGCACTTTGGAAGAACTTTTAAAGAGCTTAAAGAGCTTAAAGAGCTTAAAGAGCTTAAAG 2039
 1961 CGGGCAGTCTGTAGATC-CTCAAGAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAG 2019
 2040 TGATTTCAAAATAAAGCAGTGAAGCAATCTTTGTGTACA 2079
 2020 CGATTTCAAAATAAATAAATAAATCTACACGTTTAAATA 2059

RESULT 8
 BC026626
 LOCUS BC026626 1882 bp mRNA linear ROD 07-AUG-2002
 DEFINITION Mus musculus, clone IMAGE:4982169, mRNA, partial cds.
 ACCESSION BC026626
 VERSION BC026626.1 GI:20071837
 KEYWORDS house mouse.

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1882)
 REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK
 COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www-shgc.stanford.edu
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAC Plate: 59 Row: m Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES
 source

1. 1882
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="IMAGE:4982169"
 /cissue_type="Colon, normal. 5 month old male mouse."
 /cissue_lib="NCI CGAP_C024"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 <1..1488
 /codon_start=1
 /product="Unknown (protein for IMAGE:4982169)"
 /protein_id="AAH26626.1"
 /db_xref="GI:20071838"

CDS

BASE COUNT 644 a 339 c 372 g 527 t

Query Match 67.6%; Score 1452.2; DB 10; Length 1882;
 Best Local Similarity 88.6%; Pred. No. 0;
 Matches 1633; Conservative 0; Mismatches 198; Indels 12; Gaps 5;

QY 247 CAAGATGGTCTTCTCAGTTCACGAGAGCTCATTCGCAAGAGCTAGAATTTCTGTGAAG 306
 DB 1 CAAGATGGTCTTCTCAGTTCACGAGAGCTCATTCGCAAGAGCTAGAATTTCTGTGAAG 60
 QY 307 AAGGAACCTAGAAAAAATACTACACACATCATCATGATTTGACACACCAAAAAA 366
 DB 61 AAGGAATAGAAAAAATACTTACACACAGCGCTCACATGAGTTTGACACACTAAAAA 120
 QY 367 GACCTGGATGATTTTCGGAAGCTATTTTCATAGATTTTTCGCAAGAGAGGCGCTTCTGTG 426
 DB 121 GATCTTGATGATTTTCGGAAGCTATTTTCATAGATTTTTCGCAAGAGAGGCGCTTCTGTG 180
 QY 427 GATTGGGAAAAAATCCAGAGAGCCCTCGAAGATTCGATTCACCTATGAAAAAGATAAG 486

Db 181 GACTGGGTAAATCCAGAGACCTCCAGAAGATTCGATTCAACCCCTATGAAAGATAAAA 240
Qy 487 GCCAGGCGCTTGCCTGATAATATCTTCCGTGTTGAAACAACTAGTGGTGGTGAACCTC 546
Db 241 GCCAGAGCGCTTGCCTGATACATATCTTCTGTGTTGAAACAAGTGGTGGTGAACCTC 300
Qy 547 AATGGTGGTTGGGAACACGATGGCTGCAAGGCCCTTAAAGTCTGATTTGGTGTGAGG 606
Db 301 AATGGTGGTTGGGAACACGATGGCTGCAAGGCCCTTAAAGTCTGATTTGGTGTGAGA 360
Qy 607 AATGAGATACCTTCTGGATCTGACTGTTGAGCAAAATGACATTTGAAATAAACCTTAC 666
Db 361 AATGAGAACACCTTTTGGATCTAATCTGTTGAGCAAAATGAACTTTGAAACAAACCTTAT 420
Qy 667 AATACAGATGTTCTCTTCTTCTTAAATGAATCTTTTAAACGGATGAAGATACCAAAAA 726
Db 421 AATACAGATGTTCTCTCTCTCTTAAATGAATCTTTTAAACGGATGAAGATACCAAAAA 480
Qy 727 ATACTACAGAAGTACAATCATTTGTCGTGTAAGAAATCTACATTTCAATCAAGCAGGTAC 786
Db 481 ATACTCCAGAAGTACAATCACTGTCGTGTAAGAAATCTACATTTCAATCAAGCAGGTAC 540
Qy 787 CCGAGGATTAATGAAGTCTTACTTCTGTCGTAAGAGGAGGTCTTACTCAGGGGAA 846
Db 541 CCGAGGATCAATAAGGAATCTTACTTCTTATAGCAAAAGGATGTCTTATTTCAAGGGAA 600
Qy 847 AATACAGAAGCTTGTGTACCTCCAGTCTCATGTGATATTTTACGCCAGTCTTCTACAACTCT 906
Db 601 AATACTGAAGCTTGTGTACCTCCAGGATGAGATATCTATGCTAGTTTCTACAAATCT 660
Qy 907 GGATTTGCTGATACCTTTATAGGAGGCAAGAGTATATTTTGTGCTTAAACATAGAT 966
Db 661 GGCTTGCTCGATACCTTTATAGAGGCAAGAGTATATTTTGTGCTTAAACATAGAT 720
Qy 967 AATCTGGGTGCCAGTGGATCTGTATTTCTTAACTCATCTAATGAACCCACCCCAATGGA 1026
Db 721 AATCTGGGTGCCAGTGGATCTTTATATTTCTTAACTCATCTGATGAACCCACCCCAATGGG 780
Qy 1027 AAACGCTGTGAATTTGTCTGAAGTCAAAATAAACAACGTCGAGATGTAAGGGCGGG 1086
Db 781 AAGCGCTGTGAGTTGTCTGAGTGTCAAAACAACAACGAGAGATGTTAAGGCTGG 840
Qy 1087 ACACCTCACTCAATATGAAGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCCAAAAGCA 1146
Db 841 ACACCTCACTAGTATGAAGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCCAAAAGCA 900
Qy 1147 CATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAATATTTAATACAAAACCTATGG 1206
Db 901 CACGTTGATGAGTTCAAGTCTGTGTCAAAATTTAAATATTTTAAACAAACCAACCTATGG 960
Qy 1207 ATTTCTCTTGACGAGTTTAAAGACTGACGAGCAAAATGCCATTGACATGGAATCAAT 1266
Db 961 ATTTCTCTTGAGCAGTTTAAAGACTTCAAGAACAGATGCCATTGACATGGAATCAAT 1020
Qy 1267 GTGAATCAAAAGACTTTGGATGGAGCCCTGATGTCAATTAATAGAACTGCAGTAGGG 1326
Db 1021 GTGAATCAAAAGACTTGGATGGAGCCCTGATGTCAATTAATAGAACTGCAGTTGGA 1080
Qy 1327 GCTGCCATCAAAAGTTTGAAGTTCTCTAGTATTAATGTGCCAAGAGCGCTTTCTG 1386
Db 1081 GCTGCAATTAAGAGTTTCCAGAAATCTTTAGGATTAATGTTCGAAGGAGTGTCTTCTG 1140
Qy 1387 CCTGTCAAAACCAATCATGATCTTCTGCTGGTGTGATGTCAAACCTCTATAGTCTTAATGCA 1446
Db 1141 CCTGTCAAAACCAATCATGATCTTCTGCTGGTGTGATGTCAAACCTCTATAGCCTTAACGA 1200
Qy 1447 GGATCTCTGCAATGAGTGAAGCGGGAATTTCTACAGTGCCTCTGGTTAAATAGGC 1506
Db 1201 GGATCTCTGACCATGAGTGAAGCGTGAATTTCCCTACAGTGCCTCTGGTTAAATAGGC 1260
Qy 1507 AGTTCTTTTACGAAGGTTCAAGATTTATCTAAGAAGATTTGAAGATATACAGATATGCTT 1566

Db 1261 AGTTCTTTTACCAAGGTTCAAGATTTATCTAAGAAGGTTTGAAGATATACCGATATGCTG 1320
Qy 1567 GAATTTGGATCACCTCACAGATTTTCCAGGAGATGTGACATTTGGAAGAAATGTTTCAATTAAG 1626
Db 1321 GAATCGGACCACTCACTGTTTCCAGGAGATGTGACATTTGGAAGAAATGTTTCACTTAAG 1380
Qy 1627 GGAACGGTTATCATTTGCAAAATCATGGTGACAGAAATGTATATCCCACTCGGAGCAGTA 1686
Db 1381 GGAACAGTTATTTATTTGCAAAATCATGGTGACAGAAATGTATATCCCTCCCGAGGAGCGTG 1440
Qy 1687 TTAGAGAACAGATGTTGCTCTGGAAACCTTCCGCATCTTGGACCACTGAAATGAAAAATAC 1746
Db 1441 TTAGAGAACAGATGTTATCTGGAAACCTTCCGCATCTTGGATCACTGAATAAGCACTTT 1500
Qy 1747 TGTGGACACTT-----AAATAATGGGCT----AGTTTCTTACAATGAATGTTCTCTAGG 1797
Db 1501 CAGATACACTTCTACTAATTTATGGCTTAAAGATTTCTTATAATAAATATCTCTTAAG 1560
Qy 1798 ATTCTAAATAAGGAGGTACTTTA-CTATGTTACTGTACCTCGAGTGTGATTTTAA 1856
Db 1561 ATTATAAATAAGCAGGTACTTTATTTACTATGTTTACTCTGCACTGTGATTTTAA 1620
Qy 1857 ATAGAGTTTCTGTCAGTATGCTTTTACTTAAAGAAAGCAGATGGAGCAATCTTTCC 1916
Db 1621 A-GTAGTTTTCTGCAATGAGCTTTTACTTAAAGAAAGCATAGATGAAGCAATCTTTCC 1679
Qy 1917 TTCTTTGAAGAGATCCCAAAAGTTAGTTCTATTTAAAGTCAATATTTGTTAACTTTAA 1976
Db 1680 TTCTTTGATGAAGCCCTAGAAATTTAGTTCTTCTTTAAAGTCAATATTTACTTAACCTCAG 1739
Qy 1977 AACTGGGCACTTTGGAGAACTTTTAAACGAAGCTCAATGATGATCAGTTTGAATTCG 2036
Db 1740 AATCGGCGCAGTTCTGCTAGATC-CTCAAGAGAAAGCTTCAGTGACTGCTCCCTTTGAATTGC 1798
Qy 2037 TTGTGATTTTCAAAATAAAGCAGTGAAGCAATCTTGTGTACA 2079
Db 1799 CTCCGATTTCAAAATAAATAAATAAATCTACACGTTTACAAA 1841

RESULT 9
AF424698
LOCUS AF424698 2026 bp mRNA linear ROD 30-OCT-2001
DEFINITION Mus musculus uridindiphosphoglucosepyrophosphorylase 2 (Ugp2) mRNA,
complete cds.
ACCESSION AF424698
VERSION AF424698.1 GI:16518391
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE
AUTHORS Fuchs,S., Resch,K., Thiel,C., Platzer,M., Jockusch,H. and
Schmitt-John,I.
TITLE An inversion between mouse and man on chromosome 11/2p13-15 detected
by high-resolution comparative mapping
JOURNAL Unpublished
REFERENCE
AUTHORS Fuchs,S., Resch,K., Thiel,C., Platzer,M., Jockusch,H. and
Schmitt-John,I.
TITLE Direct Submission
JOURNAL Submitted (26-SEP-2001) Developmental Biology and Molecular
Pathology, University of Bielefeld, Universitaetsstrasse 25,
Bielefeld, NRW 33615, Germany
FEATURES
source Location/Qualifiers
1..2026
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="11"
1..2026
/gene="Ugp2"
273..1799
CDS

```

/gene="Ugp2"
/locus="Ugp2"
/codon_start=1
/product="uridine diphosphoglucose pyrophosphorylase 2"
/protein_id="AAL24807.1"
/db_xref="GI:16518392"
/translation="MSRRFVQDLSKMSODGASQFOEVLQBLELSVKKELEKILTTAA
SHEPHTKKDLGGPKLPHRLEQKGPSPDNGKIQRPEDSIOPEYKIKARGLPDNIS
SVLNKLVVKNLGGGLTSMGCKSPKSLIGVRNENFLDLTVQIEHLNKTINTDPLV
LMSFNFTDFTKKILQKNHCRVKIYFNQSRYPINKESLLPIAKDVSYSNGENTEA
YPPGHGDIYASFGSLDLDTIEBKEYIFVSNIDNLGATVDLYLHNLMPNGKRC
EFVMEVNTKADVKGGLTQVEKRLVETAKQVPHKAVDEPKFSKFKIFNTNNLWI
SLGAVKRLQEQNAIDMEIIVNPKTLDGLNVLQLETAAGAAIKSFENSLGIVNPRSF
LPVKTSDLLVMSNLSLNLGSLTMSKREPPVPLVKLGSSFTKVQDYLRRESIP
MLELHDHUTVSGDVTFGKNSLKVGTIIIAHNGRIDIPPGAVLENKIVSGLRILDDH
"
BASE COUNT      642 a   376 c   419 g   589 t
ORIGIN
Query Match      66.0%; Score 1416.2; DB 10; Length 2026;
Best Local Similarity 90.0%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 163; Indels 11; Gaps 4;
QY 224 AAGATCTTAGCAAGCAATGTCTCAAGATGGTCTTCTCAGTTCACGAAGTCAATTCGGC 283
DB 289 AAGATCTTAGCAAGCAATGTCTCAAGATGGTCTTCTCAGTTCACGAAGTCAATTCCTC 348
QY 284 AAGACTAGAAATTTATCTGTGAAGAGGAACCTAGAAAAAATCTACACAGCATCATCAC 343
DB 349 AAGAACTAGAAATTTATCTGTGAAGAGGAATTTAGAAAAAATCTTACCACAGCAGCTCAC 408
QY 344 ATGAATTTGACACACCAAAAGACCTCGATGGATTTTCGGAAGCTATTTTCATAGATTTT 403
DB 409 ATGAGTTTGACACACTAAAAGATCTTGATGGATTTTCGGAAGCTATTTTCATAGATTTT 468
QY 404 TGCAAGAAAAAGGGGCTTCTGTGGATTTGGGAAAAAATCCAGAGACCCCTCGAAGATTCGA 463
DB 469 TGCAGAAAAAGGACCGTCTGTGGATTTGGGTAATAATCCAGAGACCTCCAGAGATTCGA 528
QY 464 TTCAACCCCTATGAAAGATAAAGGACGAGGGTTCGCTGATATATATATCTTCCGTGTTGA 523
DB 529 TTCAACCCCTATGAAAGATAAAGGACGAGGGTTCGCTGATATATATATCTTCTGTGTTGA 588
QY 524 ACANAAGTGTGTGTAACCTCAATGTGTGTTTCGGAACACGATGGCTGCAAGGCC 583
DB 589 ACNAAGTGTGTGTAACCTCAATGTGTGTTTCGGAACACGATGGCTGCAAGGCC 648
QY 584 CTAAAAGTCTGATTTGTTGAGGAATGAGAATACCTTTCTGATCTGACTGTTCAGCAAA 643
DB 649 CTAAAAGTCTGATTTGTTGAGGAATGAGAATACCTTTCTGATCTGACTGTTCAGCAAA 708
QY 644 TTGAACATTTGAAATTAACCTTCAATACAGATGTTCTCTGTTTAAATGAACTCTTTTA 703
DB 709 TTGAACATTTGAAATTAACCTTCAATACAGATGTTCTCTGTTTAAATGAACTCTTTTA 768
QY 704 ACACGGATGAAGATACCAAAAAAATCTACAGAGTACAACTCAATGTCGTGTGAAATCT 763
DB 769 ACACGGATGAAGATACCAAAAAAATCTACAGAGTACAACTCAATGTCGTGTGAAATCT 828
QY 764 ACATCTTTCAATCAAAAGCAGGTACCCGAGGATTAATAAGAAATCTTTTACTTCTGTAGCAA 823
DB 829 ACACCTTTCAATCAAAAGCAGGTACCCGAGGATCAATAAGAAATCTCTACTTCTATAGCAA 888
QY 824 AGGAGTGTCTTACTCAGGGGAAAAATACAGAGCTTGGTACCTCCAGGTCAATGTGATA 883
DB 889 AGGAGTGTCTTACTCAGGGGAAAAATCTCAGAGCTTGGTACCTCCAGGGCATGGAGATA 948
QY 884 TTTAGCCAGTGTCTTACTCAGGTCTGATTTGATACCTTTTATAGGAAGCAAGAGT 943
DB 949 TCTATGCTAGTTCTTACATTTCTGGCTTCTCGATACCTTTTATAGGAAGCAAGAGT 1008
QY 944 ATATTTTGTGTCTAACATAGATAATCTGGGTGCCACAGTGGATCTGTATTTCTTAATC 1003

```

RESULT 10
 AF004368
 LOCUS Cricetulus griseus UDP-glucose pyrophosphorylase mRNA, complete cds.
 DEFINITION
 ACCESSION AF004368

```

1009 ATATTTTGTGTCTAAACATAGATAATCTGGGTGCCACAGTGGATCTTTATATTTCTTAATC 1068
QY 1004 ATCTAATGNAACCCACCCAAATCGAAACGCTGTGTAATTTGTCTCATGAGTGCACAAATAAAA 1063
DB 1069 ACCTGATGAACCCACCCAAATCGGAAGCGCTGTGAGTTTGTCTATGAAAGTGCACAAACAAA 1128
QY 1064 CACGTGCAGATGTAAAGGGCGGACACCTCACTCAATATGAAGGCAAACTGAGACTGGTGG 1123
DB 1129 CACGAGCAGATGTAAAGGTGGGACACCTCACTCAGTATGAAGGCAAACTGAGACTGGTGG 1188
QY 1124 AAATTTGCTCAAGTCCCAAAAGCACATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAA 1183
DB 1189 AAATCGCTCAAGTCCCAAAAGCACACGTTGATGAGTTCAAGTCTGTGTCAAAATTTAAAA 1248
QY 1184 TATTTAATACAAAACAACTATGGATTTCTCTTTCGAGCAGTTTAAAGACTGCGAGGACAAA 1243
DB 1249 TATTTAACAAAACAACTATGGATTTCTCTTTCGAGCAGTTTAAAGACTTTCAGGAACAGA 1308
QY 1244 ATGCCATTGACATGGAATAATCATTTGTGAATGCAAAAGACTTTTGGATGGAGCCCTGAATGTCA 1303
DB 1309 ATGCCATTGACATGGAATAATCATTTGTGAATGCAAAAGACTTTTGGATGGAGCCCTGAATGTAA 1368
QY 1304 TTCAATTTAGAAAATGTCAGTAGGGCTGCCATCAAAAAGTTTTCGAGAAATTTCTCTAGGTATTA 1363
DB 1369 TTCAGTTAGAAAATGTCAGTAGGGCTGCCATCAAAAAGTTTTCGAGAAATTTCTCTAGGTATTA 1428
QY 1364 ATGTGCCAAGGAGCGGTTTCTGCTGTCGCAAAACACATCAGATCTCTTGTGTTGATGT 1423
DB 1429 ATGTTCGAAGGAGTGGTTTCTGCTGTCGCAAAACACATCAGATCTCTTGTGTTGATGT 1488
QY 1424 CAAACCTCTATAGTCTTAAATGACAGATCTCTGCAAAATGAGTGAAGGCGGAAATTTTCCTA 1483
DB 1489 CAAACCTCTATAGTCTTAAATGACAGATCTCTGCAAAATGAGTGAAGGCGTGAATTTCCCTA 1548
QY 1484 CAGTGCCTTCTGTTTAAATTAGGCAAGTCTTTTTCGAAGGTTCAAGATTTATCTAAGAGAT 1543
DB 1549 CAGTGCCTTCTGTTTAAATTAGGCAAGTCTTTTTCGAAGGTTCAAGATTTATCTAAGAGAT 1608
QY 1544 TTGAAGTATACCAAGATATGCTTGAATGGAATCACCTCACAGTTTCAGGAGATGTGACAT 1603
DB 1609 TTGAAGTATACCAAGATATGCTTGAATGGAATCACCTCACAGTTTCAGGAGATGTGACAT 1668
QY 1604 TTGAAAAAATGTTTTCATTAAGGGAACCGTTCATCATTTGCAAAATCATGGTGCACAGAA 1663
DB 1669 TTGAAAAAATGTTTTCATTAAGGGAACCGTTCATCATTTGCAAAATCATGGTGCACAGAA 1728
QY 1664 TTGATATCCCACTGGAGCAGTATTAAGAGAACAGATTTGTCTCGGAAACCTTCCCATCT 1723
DB 1729 TTGACATCCCCCAGGAGCGGTGTAGAGAACAGATTTGTATCTGGAACACCTTCCGATCC 1788
QY 1724 TGGACCACTGAAATGMAAAATACTCTGGACACTT-----AAATATGGCT---AGTTT 1774
DB 1789 TGGATCATGAAATGAAGCACTTTCAGATACACTTTCTACTAATATGGGCTAAAGAGTTT 1848
QY 1775 CTCAAAATGAAATGTTCTCTAGGATTTTAAATAGGCAAGTACTTTA-CTATGTTACTGT 1833
DB 1849 CTTAATAAATAATCTCTTAAGATTATAAATAAGCAGGTACTTTATTTACTATGTTGT 1908
QY 1834 ACCCTGCAAGTGTGATTTTAAATAGAGTTTTCGCAAGTATGCTTTTAGTCTTAAGAAA 1893
DB 1909 ACTCTGCAAGTGTGATTTTAAAA-GTAGTTTTTCTGCAATGAGCTTTTAGTCTAAGGAAA 1967
QY 1894 GCACAGATGGAGCAATCTTCTTCTTTGAGAGAAATCCCAAAAGTTAGTTTCATCTTA 1952
DB 1968 GCATAGATGAAGCAATCTTCTTCTTCTTCGATGAAAGCCTAGAAATTAGTTTCTCTTA 2026

```

AF004368.1 GI:2358274

Cricketulus griseus.
Cricketulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.

1 (bases 1 to 1880)
Peng, H.L. and Chang, H.Y.
Cloning of a human liver UDP-glucose pyrophosphorylase cDNA by
complementation of the bacterial *galU* mutation
FEBS Lett. 329 (1-2), 153-158 (1993)

JOURNAL
MEDLINE
93359039
PUBMED

2 (bases 1 to 1880)
Flores-Diaz, M., Alape-Giron, A., Persson, B., Pollesello, P., Moos, M.,
von Eichel-Streiber, C., Thelestam, M. and Florin, I.
Cellular UDP-glucose deficiency caused by a single point mutation
in the UDP-glucose pyrophosphorylase gene
J. Biol. Chem. 272 (38), 23784-23791 (1997)

JOURNAL
MEDLINE
9742447
PUBMED

3 (bases 1 to 1880)
Flores-Diaz, M., Alape-Giron, A. and Thelestam, M.
Direct Submission
Submitted (15-MAY-1997) Microbiology and Tumorbiology Center,
Karolinska Institute, Box 280, Stockholm S-171 77, Sweden
Location/Qualifiers

1. .1880
/organism="Cricketulus griseus"
/strain="ATCC CCL-16"
/db_xref="taxon:10029"
/issue_type="lung pseudodiploid Chinese hamster Don
fibroblast"
88. .1614
/EC_number="2.7.7.9"
/function="synthesis of UDP-glucose from
glucose-1-phosphate"
/codon_start=1
/product="UDP-glucose pyrophosphorylase"
/protein_id="AAC53343.1"
/db_xref="GI:2358275"
/translation="MSRPFQDLISKAMSDGASQFQEVILQLELSVKELEKILTTAT
SHEYETHKADLQFRLVHRFLQERGPSVDMGKIQRPEDSIQYERIKARGLPDNIS
SVLNKLVVDKNGGSLTGMCKPKSLIGVRNENTFLDLTVQOIEHLNKSNTVDPLV
LWNSLFTDTTKLIIQKHCKVRIYTFNQSRPIKINSESLLPVAKOVSSGSETEAW
YPPGHDIYASFNQDGLDITFLEEGKEIVFVSDINLGATVDLYILANHLMPNGKCE
EFPMVEVTKNRADVGGLTTOYEGKRLVEIAQVPAHVDEPKSVSEKIFNTNNLMI
SLAAVKRLQEQNAIDMEIIVNPKTLDDGLNVIQLETAVGAARKSFENLVNRSF
LPVKTSLLLVMSNLYSLNAGSLTMSSEKRFPPVPLKLGSSFTKVDYLRFPESIP
DMLELDHLTVSGDVFGRKNSVLSKGVTVIIIANHGRDIPPGAVLENKIVSNLRILDH
"

BASE COUNT 617 a 334 c 395 g 534 t

ORIGIN

Query Match 65.9%; Score 1415; DB 10; Length 1880;
Best Local Similarity 88.6%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 180; Indels 25; Gaps 4;

220 GAGGAAGACTTTAGCAAGCAATGCTCAAGATGGTGTCTTCTCAGTTCCAGAAGAGTCAAT 279
100 GTGCAGACCTTAGCAAGCTATGCTCAAGATGGTGTCTTCTCAGTTCCAGAAGAGTCAAT 159
280 CGCGAAGAGCTAGAAATTTCTGTGAAGAGGAAGCACTAGAAAAAATACTCACCACAGCATCA 339
160 CTCCAGAACTAGAAATTTCTGTGAAGAAAGAAATTTAGAAAAAATACTTAGCAAGCAACC 219
340 TCACATGAATTTGACGACACCAAAAAAGACCTGGATGATTCGGAAGCTATTTTCATAGA 399
220 TCACATGAGTATGAGCATACTAAAAAAGATCTTGAATGATTTCCGGAAGCTATATCATAGA 279
400 TTTTTCGAGAAAGGGGCTTCTCGATTTGGGAAAAAATCCAGAGACCCCTGAAGAT 459

1540 AGATTGTTGAAAGTATACAGATATGCTTGATTTGGATCCTCCTACAGTTTTCAGGAGATGTG 1599
1420 AGTTTGAAGTATACCTGACATGCTGGAATTTGGATCCTCCTCCTGTTTCTGGAGATGTG 1479
1600 ACATTGTTGAAAGTATGTTTTCATTAAAGGGAACAGTTATCATATTGCAAAATCATGTTGAC 1659
1480 ACATTGTTGAAAGTATGTTTTCATTAAAGGGAACAGTTATCATATTGCAAAATCATGTTGAC 1539
1660 AGAATTGATATCCCACTGGAGCAGTATTAAGAGCAAGATTTGTTGTTGAAACCTTGGC 1719
1540 AGAATTGATATCCCACTGGAGCAGTATTAAGAGCAAGATTTGTTGTTGAAACCTTGGC 1599
1720 ATCTTGGACCACTGAATTAAGAAATCTGTCGACACTT-----AAATATGGCT---A 1770
1600 ATCTTGGATCACTAGATTAAGCGTCCGCGATACACTTTATCTAATATGTTGTTAAACA 1659
1771 GTTCTTCAATCAATGTTCTCTAGGATTTCTAAATAGGCGAGTACTTTTACTATGTTAC 1830
1660 GTTCTTGAATTAAGTCTCTAAGATTTCTCAATGAGCAAGTACTTT----- 1709
1831 TGACCTCTGAGTGTGATTTTAAATAGAGTTTCTGCGAGTATGCTTTTAAAGTCTAAGA 1890
1710 --TACTGTGAGTGTGATTTT---TAGAGTTTCTGCAATGCTTTTAAAGTCTAAG 1763
1891 AAGACACATGAGCAATCTTCTCTCTTCTTGAAGAGATCCCAAGAGTTAGTTTCACT 1950
1764 AAGAGATAGTGAAGCAGTACTGCTCTTTTGTGATGAAGAGCTAGAGATGAGTTCACT 1823
1951 TAAAGTGAATATTTGTTTAAATCTTAAACTGGGCACTTTGGAAGAACTTTTAA 2003
1824 AAGAGTGAATATTTTAACTGTAAGTCTGGCCAGCTCGCAATCTTTTAA 1876

RESULT 11
BOVUDPGPYR 1689 bp mRNA linear MAM 05-JAN-1994
LOCUS Bovine UDP-glucose pyrophosphorylase mRNA, complete cds.
DEFINITION L14019
ACCESSION L14019.1 GI:289447
VERSION UDP-glucose pyrophosphorylase.
KEYWORDS Bos taurus (strain Danshaku-Imo) (library: lmabda gtl1) liver cdna
SOURCE to mRNA.

ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1689)
Konishi, Y., Tanizawa, K., Muroya, S. and Fukui, T.
Molecular cloning, nucleotide sequencing, and affinity labeling of
bovine liver UDP-glucose pyrophosphorylase
J. Biochem. 114 (1), 61-68 (1993)
94012583
PUBMED
8407878

Location/Qualifiers
1. .1689
/organism="Bos taurus"
/strain="Danshaku-Imo"
/db_xref="taxon:9913"
/tissue_type="liver"
/tissue_lib="lmabda gtl1"
115. .1641
/EC_number="2.7.7.9"
/codon_start=1
/product="UDP-glucose pyrophosphorylase"
/protein_id="AAA30801.1"
/db_xref="GI:289448"
/translation="MSRFVQDLKAMSODGASQFQVIRQELSLVKKEKILITLTP
SHEFETKDLGFRKFLRFLQKGPVSDWKGKIQRPEDSTQPEKIKARGLPDVNS
SVNLKLVVNLKGLGTSWGCKGPKSLIGVRNENTFLDLTVQOIEHLNKTIDTDFLV
LMSNTDSTDKILQKYNHCRKIKYFNQSRYPINKESLLPFAKNYSYSGENTFAW
YPPGHGDIYASFNYSGLLDTFIEGKEYIFVSNIDNLGATVDLYTLNHLNMPNPKPC
EFWMEVNTNKRADVKGGTLTQYEGKRLVEIAQVPAKHVDFEKSFKPKIFNTNLLI

source

CDS

BASE COUNT 550 a 318 c 362 g 459 t
ORIGIN
Query Match 63.9%; Score 1372.8; DB 4; Length 1689;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1434; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 224 AAGATCTTAGCAAGCAATGTCTCAAGATGTGCTTCTCAGTTCCAAAGAGTCAATCGGC 283
DB 131 AAGATCTCAGCAAGCAATGTCTCAGGATGTGCTTCTCAGTTCCAAAGAGTCAATCGAC 190
QY 284 AAGAGCTAGATTTATCTGTGAAGAGGAACTAGAAAAATACTCACCACGATCATCAC 343
DB 191 AAGAACTAGAAATTTCTGTGAAGAGGAACTAGAAAAATACTCACCACGATCATCAC 250
QY 344 ATGAATTTGACACACCAAAAAAGACCTGGATGGATTTTCGGAAGCTATTTTCATAGATTTT 403
DB 251 ATGAATTTGAACACACACTAAAAAGACCTTTGATGGATTTTCGGAAGCTATTTTCACAGATTTT 310
QY 404 TGCAAGAAAAAGGGGCTTCTCTGTGATTTGGGAAAAAATCCAGAGACCCCTCGAAGATTCGA 463
DB 311 TGCAAGAAAAAGGGGCTTCTCTGTGATTTGGGAAAAAATCCAGAGACCTCGGAAAGATTCGA 370
QY 464 TTCAACCTTATAAAAAGATAAAGGCCAGGGGCTTCCCTGATATAATATATCTCCGTTTGA 523
DB 371 TTCAAGCCCTATGAGAAGATAAAGGCCAGGGGCTTACCTGATGATGATATCTTCTGTGCTGA 430
QY 524 ACAAACTAGTGTGTGAAACTCAATGTGTGTTTGGGAAACCAAGCATGGCTGCAAGGCC 583
DB 431 ACAAGCTGTGTGTGAAACTCAATGTGTGTTTGGGAAACCAAGCATGGCTGCAAGGCC 490
QY 584 CTAAAGTCTGATTTGTGTGAGGAATAGAAATACCTTTCTGGATCTGACTGTTTCAGCAAA 643
DB 491 CAAAGAGTCTGATTTGGCGTGAGGAACGAGAAATACCTTTCTGGATCTGACCGTTTCAGCAA 550
QY 644 TTGAACATTTGAATAAAACTTACATACAGATGTTCTCTCTGTTTAAATGAATCTTTTA 703
DB 551 TCGAACTTTGAACAAAAACCTACGATACGATGTTCTCTCTGTTTAAATGAATCTTTTA 610
QY 704 ACACGGATGAAGATACCAAAAAATATCTACAGAAGTACAATCATTTCTGCTGTGAAAAATCT 763
DB 611 ACACGGATGAAGATACCAAAAAATATCTACAGAAGTACAATCATTTCTGCTGTGAAAAATCT 670
QY 764 ACATTTTCAATCAAAAGCAGGTACCCGAGGATTAATAAGAAATCTTTTACTTCTGTAGCAA 823
DB 671 ACACGTTTAAATCAAAAGCAGGTACCCAGGATTAATAAGAAATCTGTTACTGCTGTAGCAA 730
QY 824 AGGAGCTGTCTTACTCAGGGGAAATACAGAAGCTTTGGTACCCCTCCAGGTCATGTTGATA 883
DB 731 AGAATGTATCATCTACTCAGGGGAAATACAGAAGCTTTGGTATCTCTCCGGGTCTAGGAGATA 790
QY 884 TTTACGCCAGTTTCTACAACTCTGATTTGCTGATACCTTTTATAGGAGAAAGCAAGAGT 943
DB 791 TTTATGCCAGTTTCTACAACTCTGATTTGCTGACACCTTTTATAGGAGAAAGCAAGAGT 850
QY 944 ATATTTTGTGTCTAACATAGATAAATCTGGGTGCCACAGTGGATCTGATATTTCTTAATC 1003
DB 851 ATATTTTGTGTCTAACATAGATAAATCTGGGTGCCACAGTGGATCTTATATTTCTTAATC 910
QY 1004 ATCTAATGAACCCCAATGGAAAAACCTGTGAATTTTGTCTAGGAGTCACAAAATAAA 1063
DB 911 ATCTCATGAACCCCAATGGAAAAACCTTTGTAATTTTGTCTAGGAGTCACAAAATAAA 970
QY 1064 CAGGTGCAGATGTAAGGGGCGGACACTCAGTCAATATAGAGGCAAACTGAGACTGGTGG 1123
DB 971 CCGGTGCCAGTGAAGGGTGGGACACTCAGTCAATATAGAGGCAAACTGAGACTGGTGG 1030
QY 1124 AAATGCTCAAGTGCACAAAGCAATGTAGACGAGTTCAAGTCTCTGTATCAAAAGTCAAAA 1183

Db 1031 AATTGCTCAAGTGCACAAAGGCGCATGTTGATGAGTTCAAAATCTGTATCAAAATCAAAA 1090
Qy 1184 TATTTAATCAAAACACCTATGATTTCTCTTGCAGCAGTAAAGACATGACGAGCAAA 1243
Db 1091 TATTTAACCAAAACACCTATGATCTCTTCTGCTGCACTTAAAGACTGACGAGGAGA 1150
Qy 1244 ATGCCATTGACATGGAATCATTTGAAATGCAAGACTTTGATGAGGCGCTGAATGTCA 1303
Db 1151 ATGCTATTGACATGGAATCATTTGAAATGCAAGACTTTGATGAGGCGCTGAATGTGA 1210
Qy 1304 TTCAATTAGAAATGCTGAGTGGGCTGCCATCAAAAGTTTTCAGAAATCTCTAGGTATTA 1363
Db 1211 TTCAGTTAGAAATGCTGAGTGGGCTGCCATTAAGTTTTCAGAAATCTCTAGGTATTA 1270
Qy 1364 ATGTGCAAGGAGCGCTTTTTCCTGCTGTCAAAACACATCAGATCTCTTGTGCTGATGT 1423
Db 1271 ATGTTCTTAGAAGCGCTTTTTCCTGCTGTCAAAACACATCAGATCTCTTGTGCTGATGT 1330
Qy 1424 CAACCTCTATAGTCTTAATGAGGATCTCTGACATGAGTGAAGCGGGAATTTCTTA 1483
Db 1331 CAACCTCTACAGCCTTAATGAGGATCTTTGACATGAGTGAAGCGGGAATTTCTTA 1390
Qy 1484 CAGTGCCTTGTGTTAAATTAGGCGCTTTTTCAGGAGTTCAGAAATCTCTAAGAGAT 1543
Db 1391 CAGTGCCTTGTGTTAAATTAGGCGCTTTTTCAGGAGTTCAGAAATCTCTAAGAGAT 1450
Qy 1544 TTGAAAGTATACAGATATGCTTTGATGAGTCACTCACAGTTTCAGGAGATGTGACAT 1603
Db 1451 TTGAAAGTATACAGATATGCTTTGATGAGTCACTCACAGTTTCAGGAGATGTGACAT 1510
Qy 1604 TTGAAAGTATGTTTCAATTAAGGGAACGGTTATCATCATTCGAAATCATGTGACAGAA 1663
Db 1511 TTGGAAGAAGTGTTCATTAAGGGAACGGTTATCATCATTCGAAATCATCGGAGCGGAA 1570
Qy 1664 TTGATATCCCACTGAGCAGTATAGAGAACAGATTTGCTGTCGAAACCTTTCGATCT 1723
Db 1571 TTGATATCCCACTGAGCAGTATAGAGAACAGATTTGCTGTCGAAACCTTTCGATCT 1630
Qy 1724 TGGACCACTGAATGAAATATCTGTGGACACTTAA 1759
Db 1631 TGGACCACTGAATGAAATATCTGTACACACTTTA 1666

RESULT 12
AY069422
LOCUS 2029 bp mRNA linear INV 17-DEC-2001
DEFINITION Drosophila melanogaster LD13601 full length cDNA.
ACCESSION AY069422
VERSION AY069422.1 GI:17862179
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2029)
AUTHORS Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C.J.,
Nunoo, J., Pacleb, J., Parages, V., Park, S., Phouanavong, S., Wan, K.,
Yu, C., Lewis, S.E., Rubin, G.M. and Celniker, S.
Direct Submission
TITLE Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,
JOURNAL Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity

within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

location/Qualifiers
1..2029
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="67A9-67A9"
1..2029
/gene="CG4347"
/note="alignment with genomic scaffold AE003552"
/db_xref="FLYBASE:FBgn0035978"
130..1692
/gene="CG4347"
/note="Longest ORF"
/codon_start=1
/product="LD13601p"
/protein_id="AAL39567.1"
/db_xref="GI:17862180"
/db_xref="FLYBASE:FBgn0035978"
/translation="MPSGIRPDIIVLLERAGVRGHRAPSDSKSEFHEVTKRDLRLLEH
DVRLLLETERKAPQALKAEMGRFADLFGRFQEGEPALDNKIKQIPENAYMYSNL
KSPNEQNEIRMLDKLVIKINGLGLTSMGCKPKSVIPVRSDITFLDLTVQQTIEHL
NKTVDANPLVMNSFNTDETEKIVRYKGRFVQIHTFNOSCFPRISREHYLPVAKD
PDVKEOMEAVYPRGSDPDTFRNSGLLKKFIEGSEYCFCLNSINDLGTVDNLINK
LVGERATTPVEFVEVTDKTRADVKGKGLIOMENKRLLEIAQVPPHVDVDFKVK
PKFNTNINWANLAAIDRLRLRLNMEIIVNNKTLNGRTVRIQLETVAGAAKCFDG
AIGNVPNSREFLPVKSSDLLVMSNLYTLKNGSLVMSQPMFPPTPLVKGHNHFSK
VKFELGRFANIPDIIEHLDTVSGDVTGFRGSLRGTVIIIANHGDRIIDIPAGAILEN
KIVSGNMRILDH"
550 a 471 c 493 g 515 t

BASE COUNT 550 a 471 c 493 g 515 t
ORIGIN

Query Match 25.1%; Score 538.8; DB 3; Length 2029;
Best Local Similarity 63.7%; Pred. No. 4.3e-120;
Matches 886; Conservative 0; Mismatches 492; Indels 12; Gaps 4;

Qy 360 CAAAGAACACCTGATGATTCGGAAGCTATTTTCATAGATTTTTCGCAAGAAAGGGCC 419
Db 312 CAAAGCAAGAAATGGGTGCTTTTCGGCGATCTGTTCGGACGTTTTCATTCAGGAAGGGTCC 371
Qy 420 TTCTGTGATTGGGGAAGAAATCCAGAGACCCCTCAAGATTCGATTCAACCCCTATGAAA 479
Db 372 CGCATTCGACTGGAACAGATTCAGAAAGCTGCTGAGAAATGCTGTGATGAATCTCGNA 431
Qy 480 GATAAAGC---CAGGGGCTTGCCTGATAATATATCTTCCTGGTGTGAAACAACTAGTGT 536
Db 432 TCTTAAGTCGCCCAAGAACGAGCAATGAGATCCGCAACATGTTGGATAAATTTGGTGT 491
Qy 537 GGTGAATCAATGATGTTGGGAACAGCATGCGCTGCAAGGCCCTCAAAAGTCTGAT 596
Db 492 CATCAAGCTGAACCGTGTGTTGGGCACCTCAATGCGTGGTCCCAAGAGTGTGAT 551
Qy 597 TGGTGTGAGGAATGAGATACCTTTCTCGATCTGATCTTCTGAGCAAAATGAACATTTGAA 656
Db 552 TCTGTGGCTGGATCTTACTTTCTGGATCTGACCGTGCAGCAAAATCGACATCTCAA 611
Qy 657 TAAAAACCTACATACAGATGTTTCTCTTCTTTTAAATGAATCTTTTAAACCGGATGAAGA 716
Db 612 CAAGACCTACGATGCCAATGTTCCGCTGCTGATGAATCTTCTTAAACACCGACGAGA 671
Qy 717 TACCAAAAAAATACCTACAGAGATCAATCATTTGTCGTGTGAAAAATCTACACTTCAATCA 776
Db 672 TACCAGAGAAGATCGTGCACAAAGTACAAAGGATTCGGTGTGACAGATCCACACCTCAACCA 731

Qy	777	AAGCAGGTACCCGAGGATTAATAAAGAAATCTTTACTTCTGTAGCAAAAGGACGTGTCTTA	836
Db	732	GAGCTGCTCCCGCGCATTAAGTCGCGAGCACTACCTGCGGTGGCCAAAGACTTTTGATGT	791
Qy	837	CTCAGGGAAAAATACAGAAGCTTGGTACCCTCCAGGTCATGGTGATATTTACGCCAGTTT	896
Db	792	CGAAAG---ATATGGAAGCCTGGTATCCACCTGGACACGGTGACTTCTACGATACCTT	848
Qy	897	CTACAACCTCTGGATTTGCTTGATACCTTTTATAGGAGAAAGCAAGAGATATATTTTGTGTC	956
Db	849	CCGCAACTCTGGCCCTGCTGAAGAAGTTTCATCGAGGAGCGCGGAGTATTGTTCTTGTGC	908
Qy	957	TAACATAGATAATCTCGGTGCCACAGTGGATCTGTATATTCTTAAATCATCTAATG---AA	1013
Db	909	CAACATCGATAAACCCTGGGCGCACCGTCGATCTGAACATCCTCAATAAGCTGGTGGCGA	968
Qy	1014	CCCACCCAATGGAAAAACGCTGTGAATTTGTTCATGGAAGTCACAATAAAAAACAAGTCGAGA	1073
Db	969	GGAGCGGCCACTACTCCAGTGGAGTTTGTCTGAGGTCACCGACAAGACCCGCTGCTGA	1028
Qy	1074	TGTAAGGGCGGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGGAAATTTGTCTCA	1133
Db	1029	CGTTAAGGGTGGTACTCTCAATTCAAATGGAAAAACAAGCTCGCGCTGTGGAAATTTGTCTCA	1088
Qy	1134	AGTSCCAAAAGCACATCTAGACGAGTTCAAAGTCGTATCAAAGTTCAAAATATTTTAATAC	1193
Db	1089	AGTGCCACCAAGCATGTGGATGACTTTAAGTCGTTAAGACCTTCAAGTTCCTTTAAACAC	1148
Qy	1194	AAACAACCTATGGATTTCTCTTGACGAGTTTAAAGACTCGAGAGCAAAATGCCATTTGA	1253
Db	1149	CAACAACATTTGGGCCAACCTCGGCAGCCATTTGATCGCGTTTGGCTGAGCGGTACTTTAAA	1208
Qy	1254	CATGGAATCATTTCTGTAATGCAAGACTTTGGATGGAGGCTGAATGTCATTCAATTTAGA	1313
Db	1209	CATGGAAATTTATTTGTGAACAACAGCGCTGGAGAACGGAACCCGCTGTCATTGACTGGA	1268
Qy	1314	AATGTCAGTAGGGCTGCCATCAAAAGTTTGTAGAATTTCTCTAGGTATTAAATGTGCCAAG	1373
Db	1269	GACTGCTGTGGGTGCGGCCATGAAGTCTTTGATGGCGCCATCGGCATCAATGTGCGCGC	1328
Qy	1374	GAGCGGTTTCTGCCTGTCAAAACCATCAGATCTCTGCTGGTGATGTCAAACCTCTA	1433
Db	1329	TTCCCGTTTCTTGCCCGTGAAGAAATCTTCAGATCTGCTGGTGATGTCAAATTTATA	1388
Qy	1434	TAGTCTTTAATGCAAGATCTCTGCAATAGTGAAAGCGGGAATTTCTTACAGTGCCTTT	1493
Db	1389	TACGCTTTAAGAACCGGACCTGGTAAATGTGCGCCACAGCGAATGTTCCCTACCAGCGCT	1448
Qy	1494	GGTTAAAT---AGGCAAGTCTTTTACGAAGGTTCAAGNTATCTAAGAAGATTTGAAG	1550
Db	1449	CGTCAAGCTGGGAGAGAACCACTTCTCGAAGGTCAAGGAGTTCCCTGGTCCGTTTGGCAA	1508
Qy	1551	TATACCAAGATATGCTTCAATTTGGATCACCTCACAGTTTTCAGGAGATGTGACATTTGAAA	1610
Db	1509	CATCCCGACATCATTGAGCTGGATCACTTGACCGTGAGCGGAGATGTGACTTCGCTCG	1568
Qy	1611	AAATGTTTCAATTAAGGGAACGGTTTATCATCTATTTGCAAAATCATGGTGACAGAATTGATAT	1670
Db	1569	TGGCGTTTCTTTGCGCGCACTGTGTATCATCTATTGCTTAACCATGGCGATCGTATCGATAT	1628
Qy	1671	CCCACTGGAGCAGTATTAGAGAACAGATTGTCTGTGGAAACCTTCGCAATCTTGGACCA	1730
Db	1629	CCCAGCTGGAGCCATTCTGGNAAACAAGATTGTATCTTGGCNATATGCGCATCTTGGACCA	1688
Qy	1731	CTGAAATGAA	1740
Db	1689	CTAATCAAA	1698

RESULT 13
AC092588
LOCUS

AC092588 140401 bp DNA linear PRI 01-MAR-2002

Homo sapiens BAC clone RP11-44A12 from 2, complete sequence.

AC092588 AC007910

VERSION AC092588.2 GI:15668104

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 140401)

Sulston,J.E. and Waterston,R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

99063792

PUBMED 9847074

2 (bases 1 to 140401)

Goyes,E., Cotton,M., Hawkins,M. and Elliott,G.

The sequence of Homo sapiens BAC clone RP11-44A12

Unpublished (2001)

3 (bases 1 to 140401)

Waterston,R.H.

Direct Submission

Submitted (19-JUL-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 140401)

Waterston,R.H.

Direct Submission

Submitted (19-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 140401)

Waterston,R.

Direct Submission

Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Sep 19, 2001 this sequence version replaced gi:14916173.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@watson.wustl.edu

----- Summary Statistics

Center project name: H_NH0044A12

Drafting Center: WIBR

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

SOURCE INFORMATION: The RPci-111 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frenken, E., Taten, M., Catanesse, J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-547M24, 2000 bp overlap.
Actual start of this clone is at base position 1 of RP11-44A12;
actual end is at base position 24890 of RP11-547M24.

. Data from AC013466 was used to finish this clone, AC007910.

The sequence of AC007910 has been incorporated into AC092588.

FEATURES

```

source
1..140401
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-44A12"
/clone_lib="RPC1-11"

misc_feature
1..175
/notes="match to EST BE180465 (NID:g8659641)"

misc_feature
1..175
/notes="match to EST BG111523 (NID:gl2605029)"

misc_feature
1..175
/notes="match to EST BG756775 (NID:gl4067428)"

misc_feature
1..175
/notes="match to EST R30886 (NID:g786729) yH59b04.r1"

misc_feature
1..175
/notes="similar to Homo sapiens EST AL580952 (NID:gl2947472)"

misc_feature
1..175
/notes="similar to Homo sapiens EST BF693345 (NID:gl1978753)"

misc_feature
1..175
/notes="similar to Mus musculus EST BI249688 (NID:gl4797311)"

misc_feature
1..175
/notes="similar to Mus musculus EST BI691328 (NID:gl15653957)"

misc_feature
32..175
/notes="similar to Mus musculus EST BE132707 (NID:g8595207) uf21g01.x1"

misc_feature
34..175
/notes="match to EST BE856601 (NID:gl0369785)"

misc_feature
157..175
/notes="match to EST BF841835 (NID:gl2195137)"

misc_feature
194..368
/notes="match to EST BF843376 (NID:gl2197958)"

repeat_region
470..808
/rpt_family="L2"

repeat_region
810..1383
/rpt_family="MER2_type"

repeat_region
1565..2115
/rpt_family="L2"

misc_feature
1916..2350
/notes="match to EST AV72588 (NID:gl0825226)"

misc_feature
1965..2462
/notes="match to EST W95984 (NID:gl1425930) ze09e01.r1"

misc_feature
2360..2716
/notes="match to EST BG621534 (NID:gl3672905)"

misc_feature
2803..2716
/notes="match to EST AA927024 (NID:g3075921) om69c06.s1"

misc_feature
2805..2716
/notes="match to EST BR076208 (NID:g8425800)"

misc_feature
2612..2717
/notes="match to EST R30886 (NID:g786729) yH59b04.r1"

misc_feature
2612..2716
/notes="match to EST BE180465 (NID:g8659641)"

misc_feature
2612..2716
/notes="match to EST BE856601 (NID:gl0369785)"

misc_feature
2612..2716
/notes="match to EST BF841835 (NID:gl2195137)"

misc_feature
2612..2716
/notes="match to EST BG111523 (NID:gl2605029)"

misc_feature
2612..2716
/notes="match to EST BG756775 (NID:gl4067428)"

misc_feature
2612..2716
/notes="match to EST R80882 (NID:g857163) yI94b10.r1"

misc_feature
2641..2716
/notes="similar to Bos taurus EST BG223793 (NID:gl2709314)"

misc_feature
2669..2716
/notes="match to EST BG289083 (NID:gl3044569)"

misc_feature
2675..2716
/notes="match to EST AA151103 (NID:gl1722652) zL48e03.s1"

misc_feature
3096..3557
/notes="match to EST AI807208 (NID:g5393774) wf38b01.x1"

misc_feature
3223..3610
/notes="match to EST R91983 (NID:g9595523) yp95f11.r1"

misc_feature
3487..3553
/notes="similar to Sus scrofa EST BF075430 (NID:gl0868941)"

misc_feature
3569..4041
/notes="similar to Sus scrofa EST BF075430 (NID:gl0868941)"

misc_feature
3626..4073
/notes="similar to Mus musculus EST AW113476 (NID:g6826189)"

misc_feature
3643..4094
/notes="similar to Mus musculus EST BI454120 (NID:gl5244776)"

misc_feature
3649..4122
/notes="match to EST AA923174 (NID:g3070483) om48a03.s1"

misc_feature
3649..4079
/notes="similar to Bos taurus EST AV615997 (NID:g9751667)"

misc_feature
3651..4125
/notes="similar to Mus musculus EST BI691328 (NID:gl15653957)"

misc_feature
3651..4096
/notes="match to EST BE856601 (NID:gl0369785)"

misc_feature
3651..4093
/notes="match to EST AA151103 (NID:gl1722652) zL48e03.s1"

misc_feature
3651..4093
/notes="match to EST BG621534 (NID:gl3672905)"

misc_feature
3651..4091

Query Match 24.4%; Score 523; DB 9; Length 140401;
Best Local Similarity 100.0%; Pred. No. 2,6e-116;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1625 AGGGAACGGTTATCATTCATTCGAAATCATGTGACAGATTGATATCCACTCGAGCAG 1684
DB 3649 AGGGAACGGTTATCATTCATTCGAAATCATGTGACAGATTGATATCCACTCGAGCAG 3708
QY 1685 TATTAGAACCAAGATTGTCTCGAAACCTTCGACCTTCGACCTGGAATGAAAT 1744
DB 3709 TATTAGAACCAAGATTGTCTCGAAACCTTCGACCTTCGACCTGGAATGAAAT 3768
QY 1745 ACTGTGACACTTAAATAATGGGTAGTTCTTACAATGAATGTTCTCTAGGATCTTAA 1804
DB 3769 ACTGTGACACTTAAATAATGGGTAGTTCTTACAATGAATGTTCTCTAGGATCTTAA 3828
QY 1805 AATAGGAGGTACTTTACTATGTACCTGCGAGTGTGATTTTAAATAGAGTT 1864
DB 3829 AATAGGAGGTACTTTACTATGTACCTGCGAGTGTGATTTTAAATAGAGTT 3888

```


ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS Bishop, J.D. and Gomer, R.H.
TITLE A second UDP-glucose pyrophosphorylase gene in Dictyostelium
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1625)
AUTHORS Bishop, J.D. and Gomer, R.H.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1999) HHMI/Biochemistry, MS-140, Howard Hughes
Medical Institute, Rice University, 6100 South Main Street,
Houston, TX 77005-1892, USA

FEATURES
source
1..1625
/organism="Dictyostelium discoideum"
/strain="Ax4"
/db_xref="taxon:44689"
/note="plasmid SSA730 supplied by the Japan Dictyostelium
discoideum cDNA project"
1..1625
/gene="UDGP2"
88..1596
/gene="UDGP2"
/EC_number="2.7.7.9"
/note="UDP-glucose-1-phosphate uridylyltransferase"
/codon_start=1
/product="UDP-glucose pyrophosphorylase 2"
/protein_id="A034028.1"
/db_xref="GI:4929335"
/translation="MKKPSLPLQSPQLSFGSRSSDLATEDFLKLEAISOTAP
NETVNEFLNKEIPINKLFLRKNKVIDMDKINPPADMLVNLKDLPAITEQRT
SELASKLAVLKLNGGLTGMCTGPKSVIEVRSEKTFDLDSVQOIKEMNRYNIKPL
VLMSNFTHQETGKIIOKYKSDVKIHSFNOSRPRILKDNLPVDPKLFPSDSWYP
PCHGVDFALONGLETLINSEKVLFTSNVDNLGAVVDNLEAMDKNVEYIMEV
TNKRDVKGGLTIOYEGKAKLLEIAQVPSKVEEFKIKPKFINTNNINVLKAMD
RILKNLDDMDIINPKVADGNIIQLEIAAGAAIEFNNGRVNVRSPFLPVKST
SLFIVQSNLSLEGLVMKNRPFITVPLVKGDNFKVSDYQARLKGIPDILELD
OLTVSGDITFGNMYLKGTVIIVANHGSRIDIPGESEFENKVVSNLHCGAL"
BASE COUNT 614 a 221 c 262 g 528 t
ORIGIN

Query Match 20.9%; Score 448.4; DB 3; Length 1625;
Best Local Similarity 61.0%; Pred. No. 4e-98;
Matches 829; Conservative 0; Mismatches 501; Indels 30; Gaps 5;

QY 361 AAAAAAGCCTGGATTCGGAGCTATTCATAGATTTTCGAGAAAAGGGCCT 420
DB 247 AACAAGAAATACCATCTATCAACAAATTTGTTTACCCGTTTCTTAAAGAAATAGAAAGAA 306
QY 421 TCTGTGGATTGGGAAAAATCAGAGACCCCTCGAAGATTCGATTCACCCCTATGAAAAG 480
DB 307 GTAATCGATTGGATAAAATCAATCCACCACGCTGATATGGTTTAAACTATAAGAT 366
QY 481 ATAAAGCCAGGGCTGCTGATATATATCTTCGGTGTGGAACAACTAGTGGTGGT 540
DB 367 TTACAGCAA---TCACTGAACAAAGAACTTCAGAACTCGCTTCAAAATTTGGCTGTTTG 423
QY 541 AACTCAATGGTGTGGTGGACCAAGCATGGCTCGAAGGCCCTTAAAGTCTGATGGT 600
DB 424 AAATTAATGGAGGTTTAGTACAACAAATGGGTGTGCTGGTCCAAAGATGTTATTGAA 483
QY 601 GTGAGGAATGAGAATACCTTTCTGGATCTGACTGTTCAGCAAAATGCAACATTTGAATAA 660
DB 484 GTTAGAAGTGAAAAAACTTTCTAGATTTAAGTTTCAACAAATTAAGAAATGATGAA 543
QY 661 ACCTACAATCAGATGTTCTCTGTGTTTAAAGTCTTTTAAACCGGATCAAGATACC 720
DB 544 AGATATAATATTAAGTACCAATAGTATTAATGAATTCATTTAATCACTCAAGAACT 603
QY 721 AAAAAAATCAAGAGTACATATCTTCGTGTGAAAAATCTACATTTCAATCAAAGC 780
DB 604 GGTAAATCATTTCAAAAAATATAAATATCTGTATGTAAAGATTCATTTCTTTAATCAAAGT 663

Search completed: June 28, 2003, 06:24:31
Job time : 3678.46 secs

QY 781 AGGTACCCAGGATTAATAAAGAAATCTTTACTTCTGTAGCAAAAGGACGTGTCTTACTCA 840
DB 664 CGTTTCCCAAGAACTTTAAAGAGATAATTAATGCCAGTACCAGA-----TAAATTA 714
QY 841 GGGGAAATACAGAAAGCTTGGTACCTCCAGTCAATGGTATATTATTCGCCAGTTTCTTAC 900
DB 715 TTTGGTAGTGATCTGAATGGTATCCACCAGTCAATGGTATGTTTCTTTGGCATTACAA 774
QY 901 AACTCTGATTTGCTGATACACCTTTATAGGAAGCAAGAGATATATTTTGTGTCTAAC 960
DB 775 AATTCTGTCTCTTGGAAACCTTATCAATGAAGGTAAAGATATTTTATTCATTTCAAT 834
QY 961 ATAGATAATCTGGGTGCGACAGTGTATATTTCTTAATCATCTAATGAACCCACCC 1020
DB 835 GTAGATAATCTGGTGTCTGTTGTTGATTTTAAACATT-----TTAGAAGCAATG 882
QY 1021 AATGGAAACCTGTGAATTTGTGATGAACTCAATTAACACGTCGAGATGAAG 1080
DB 883 GATAAAAATAAAGTTGAATATATTATGGAAGTTTCAAAATAAACTAGAGCAGATGTTAAA 942
QY 1081 GCGGGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCA 1140
DB 943 GGTGGTACATTAATCAATATGAAGTTAAAGCTAAATTTATGGAGATTCACCAAGTACA 1002
QY 1141 AAAGCACATGTAGACGAGTTTCAAGTCTGTATCAAAGTTTCAAAATATTTTAATACAAAC 1200
DB 1003 TCATCAAAGTTGAAGAAATTTAAATCAATAAAGAAATTTCAAAATTTTAAATACCAATAAC 1062
QY 1201 CTATGGATTTCTTTGACGAGTTTAAAGACTGACGAGGCAAAATGCCAT---TGACATG 1257
DB 1063 ATTTGGGTCAATCTCAAAGCAATGGATGATTTTGAACAAAAATCTTTTGGATGATATG 1122
QY 1258 GAAATCATTTGTAATGCAAGACTTTGATGGAGGCTTGAATGTCTTCAATTTAGAAACT 1317
DB 1123 CATATCATTAATCAATCAAAG---GTTGCTGATGTTAAATATCATTTCAATTTGAGAT 1179
QY 1318 CAGTATGGGTGGCATCAAAGTTTGAAGATTTCTTATAGTATTAATGTGCCAAGGAGC 1377
DB 1180 GCTGCTGGTGTGCAATTTGAATTTCTTAAATGCAAGAGGTGTCAATGTACCAAGATCA 1239
QY 1378 CGTTTCTGCTGTCAAACCAACCATCATCTCTTGTGGTATGTCACCACTCTATAGT 1437
DB 1240 AGATTTTACAGTGAAATCAACCTCTGATCTCTTTATCGTTCAATCAATCTCTACTCA 1299
QY 1438 CTTAATGCAGGATCTCTGACAAATGAGTGAAAGGCGGAATTTCTTACAGTGCCTTGGTT 1497
DB 1300 CTCGAAAGGGTGTGTTGGTAAAGCAAGAAATAGACCATTCACCAAGTTCCTTGGTT 1359
QY 1498 AAATTAGGAGTCTTTTACGAAGGTTCAAGATTTATCTAAGAAGATTTTGAAGTATACCA 1557
DB 1360 AAATTTGGGTGATAACTTTTAAAGGTTTCAGATTTATCAAGCTCGTATCAAAGGTATTCCA 1419
QY 1558 GATATGCTTGAATGGATCACCTACAGTTTTCAGAGATGTGACATTTTGGAAAAAATGTT 1617
DB 1420 GATATTTTGGAAATTTGGATCAACCTTCTGCTGATATAAACCCTTTGGTCCAAATATG 1479
QY 1618 TCATTTAAAGGAAACGGTTTATCATCTTCAAAATCATGTGTGACAGAAATTTGATCCACCT 1677
DB 1480 GTCCTCAAGGACATGTTTATCATTTGCTGCTAATCATGGTTCAAGNAATTTGATTTCTGAA 1539
QY 1678 GGAGCAGTATTAGAGAAACAAGATTTGTGCTCTGGAACCTTC 1717
DB 1540 GGTTCAGAAATTTGAAATTAAGTAGTAGTAGTGGTAAATTTTAC 1579

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:03:00 ; Search time 310.377 Seconds
(without alignments)
15577.944 Million cell updates/sec

Title: US-09-981-353-104
Perfect score: 2147
Sequence: 1 gtgttaggagaaagttagg.....gttactatattagttggaag 2147

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : 1_ /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1676.6	78.1	1823	24	ABN96841
2	1676.6	78.1	1823	24	ABL62448
3	828	38.6	853	21	AAC10871
4	675	31.4	709	22	AA197888
5	626.2	29.2	744	22	AA197887
6	622.6	29.0	650	24	ABQ58724
7	535.8	25.0	2187	23	ABL07495
8	535.8	25.0	2222	23	ABL07483
9	519.2	24.2	701	24	ABQ58840

C	10	419.8	19.6	671	24	ABK62970	Rat sequence diffe
	11	392.2	18.3	2045	21	AAA67125	Pinus radiata UGP
	12	385.4	18.0	1960	21	AAA67124	Eucalyptus grandis
C	13	375.2	17.5	396	22	AAF94852	Human ovarian can
C	14	375.2	17.5	396	24	ABT03119	Human ovarian can
C	15	375.2	17.5	396	24	ABL48802	Ovarian carcinoma
	16	374	17.4	1398	17	AAAT03700	Cotton UDP glucose
	17	368	17.1	619	21	AAC00680	Human secreted pro
	18	361.6	16.8	1814	21	AAA59677	DNA encoding a uri
	19	356.6	16.6	1700	21	AAA59676	DNA encoding a uri
	20	356	16.6	1458	21	AAC48163	Arabidopsis thalia
	21	355.6	16.6	1744	21	AAC44032	Zea mays DNA fragm
	22	355.6	16.6	1744	21	AAC44032	DNA encoding a uri
	23	350.2	16.3	487	21	AAA59672	Arabidopsis thalia
	24	349.6	16.3	1756	21	AAC39633	Human colon cancer
C	25	324.2	15.1	634	21	AAA16506	Arabidopsis thalia
	26	267.6	12.5	1262	21	AAC35937	Human colon cancer
	27	246.2	11.5	1036	21	AAA67094	Arabidopsis thalia
	28	237.2	11.0	8518	23	ABL07482	Pinus radiata UGP
	29	237.2	11.0	9585	23	ABL07482	Drosophila melanog
	30	217.6	10.1	373	14	AAQ60218	Human brain expres
	31	205.4	9.6	383	24	ABQ60743	Human colon cancer
C	32	181.8	8.5	655	24	ABQ60410	Human colon cancer
	33	170	7.9	1108	21	AAAF7092	Pinus radiata UGP
	34	166.4	7.8	773	21	AAAF7092	Rapergillus oryzae
C	35	151	7.0	467	21	AAAF7095	Eucalyptus grandis
	36	138.8	6.5	3540	23	ABL07654	Drosophila melanog
	37	128.6	6.0	798	20	AAV69481	Banana fruit ripen
	38	115.8	5.4	519	21	AAA59673	DNA encoding a uri
	39	110	5.1	464	22	ABA52276	Human foetal liver
	40	110	5.1	464	22	ABA22077	Probe #543 for gen
	41	110	5.1	464	22	AAK00551	Human brain expres
	42	110	5.1	464	22	AAK26000	Human bone marrow
	43	110	5.1	464	22	AAI10629	Probe #562 for gen
	44	110	5.1	464	22	AAI31883	Probe #569 used to
	45	110	5.1	464	22	AAI00560	Probe #551 used to

ALIGNMENTS

RESULT 1
ABN96841
ID ABN96841 standard; DNA; 1823 BP.

AC ABN96841;

DT 13-AUG-2002 (first entry)

DE Gene #3339 used to diagnose liver cancer.

Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN WO200229103-A2.

PD 11-APR-2002.

PF 02-OCT-2001; 2001WO-US30589.

PR 02-OCT-2000; 2000US-237054P.

PA (GENE-) GENE LOGIC INC.

PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
WPI; 2002-426119/45.

DR Diagnosing and detecting the progression of liver cancer,

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample
 XX Claim 1; SEQ ID NO 3339; 298pp; English.
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytotactic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1823 BP; 593 A; 337 C; 402 G; 491 T; 0 other;
 Query Match 78.1%; Score 1676.6; DB 24; Length 1823;
 Best Local Similarity 98.5%; Pred. No. 0;
 Matches 1703; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
 QY 225 AGATCTTAGCAAGCAATGCTCAAGATGGTCTCTCAAGTCCCAAGAGTCAATTCGGCA 284
 DB 93 AGATCTTAGCAAGCAATGCTCAAGATGGTCTCTCAAGTCCCAAGAGTCAATTCGGCA 152
 QY 285 AGAGCTAGATATATCTGTGAAGCAAGCACTAGAAAATACTCACCACAGCATCATACA 344
 DB 153 AGAGCTAGATATATCTGTGAAGCAAGCACTAGAAAATACTCACCACAGCATCATACA 212
 QY 345 TGAATTTGAGCACACCAAAAAAGACCTGGATGGATTTCCGAAAGCTATTTTATAGATTTT 404
 DB 213 TGAATTTGAGCACACCAAAAAAGACCTGGATGGATTTCCGAAAGCTATTTTATAGATTTT 272
 QY 405 GCAAGAAAGGGGCTCTCTGGATTTGGGAAAAATCCAGAGACCCCTCGAAGATTCGAT 464
 DB 273 GCAAGAAAGGGGCTCTCTGGATTTGGGAAAAATCCAGAGACCCCTCGAAGATTCGAT 332
 QY 465 TCAACCTATGAAGAATAAGCAAGCAAGGGCTTGGCTGATATATATCTTCCGTGTGA 524
 DB 333 TCAACCTATGAAGAATAAGCAAGCAAGGGCTTGGCTGATATATATCTTCCGTGTGA 392
 QY 525 CAAACTAGTGGTGAACCTCAATGGTGGTGGGAAACAGCATGGGCTGCAAAAGGCC 584
 DB 393 CAAACTAGTGGTGGTGAACCTCAATGGTGGTGGGAAACAGCATGGGCTGCAAAAGGCC 452
 QY 585 TAAAGTCTGATGGTGGAGGAATGAGAAATACCTTTCTGGATCTGACGTTCAGCAAT 644
 DB 453 TAAAGTCTGATGGTGGAGGAATGAGAAATACCTTTCTGGATCTGACGTTCAGCAAT 512
 QY 645 TGAACATTTGAACAAACCTCAATACAGATGCTCTCTGTTGTTTAAAGAACTCTTTAA 704
 DB 513 TGAACATTTGAACAAACCTCAATACAGATGCTCTCTGTTGTTTAAAGAACTCTTTAA 572
 QY 705 CACGGATGAAGTACCAAAAAATACTACAGAAATGATCAATCTTGTGTTGAAAAATCTA 764
 DB 573 CACGGATGAAGTACCAAAAAATACTACAGAAATGATCAATCTTGTGTTGAAAAATCTA 632
 QY 765 CACTTTCAATCAAGCAGTACCGAGGATTAATAAGAAATCTTTACTCTCTGTAGCAAA 824
 DB 633 CACTTTCAATCAAGCAGTACCGAGGATTAATAAGAAATCTTTACTCGGCTGTAGCAAA 692
 QY 825 GGACGTGCTTACTCAGGGGAAAAATACAGAACTGGTACCTCCAGGTCATGGTGATAT 884
 DB 693 GGACGTGCTTACTCAGGGGAAAAATACAGAACTGGTACCTCCAGGTCATGGTGATAT 752
 QY 885 TTACGCCAGTTTCTCAACTCTGGATTTGCTTGATACCTTTATAGGAGAGGCAAGAGTA 944

DB 753 TTACGCCAGTTTCTCAAACTCTGGATTTGCTGTATACCTTTATAGGAGAGCAAGAGTA 812
 QY 945 TATTTTGTGCTCAACATAGATAATCTGGGTGGCCACAGTGGATCTGTATATTTCTTAATCA 1004
 DB 813 TATTTTGTGCTCAACATAGATAATCTGGGTGGCCACAGTGGATCTGTATATTTCTTAATCA 872
 QY 1005 TCTAATGAACCCCAACCAATGGAACACGCTGTGAATTTGTCTATGAAGTCACAAAATAAAC 1064
 DB 873 TCTAATGAACCCCAACCAATGGAACACGCTGTGAATTTGTCTATGAAGTCACAAAATAAAC 932
 QY 1065 ACGTGCAGATGTAAGGGCGGGAACCTCACTCAATATGAAGGCAAACTGAGAGCTGGTGA 1124
 DB 933 ACGTGCAGATGTAAGGGCGGGAACCTCACTCAATATGAAGGCAAACTGAGAGCTGGTGA 992
 QY 1125 AATTGCTCAAGTGCACCAAGACATGTAGAGAGTCTCAAGTCTGTATCAAAAGTTCAAAT 1184
 DB 993 AATTGCTCAAGTGCACCAAGACATGTGTAGAGTCTCAAGTCTGTATCAAAAGTTCAAAT 1052
 QY 1185 ATTTAATACAAACCAACCTATGATTTCTTCTGACAGCACTTAAAGAGCTGCAAGGCAAAA 1244
 DB 1053 ATTTAATACAAACCAACCTATGATTTCTTCTGACAGCACTTAAAGAGCTGCAAGGCAAAA 1112
 QY 1245 TGCCATTGACATGAAATCATTTGTAATGCAAAAGCACTTTGGATGGAGCCCTGAATGTCAT 1304
 DB 1113 TGCCATTGACATGAAATCATTTGTAATGCAAAAGCACTTTGGATGGAGCCCTGAATGTCAT 1172
 QY 1305 TCAATTAGNAACTGCAGTAGGGCTGCCATCAAAAGTCTTTCAGAAATCTCTAGGTATTAA 1364
 DB 1173 TCAATTAGNAACTGCAGTAGGGCTGCCATCAAAAGTCTTTCAGAAATCTCTAGGTATTAA 1232
 QY 1365 TGTGCCAAGGAGCGTTTCTGCTGTCAAAACCAACATCAGATCTTCTGCTGGTGTATGTC 1424
 DB 1233 TGTGCCAAGGAGCGTTTCTGCTGTCAAAACCAACATCAGATCTTCTGCTGGTGTATGTC 1292
 QY 1425 AAACTCTATAGTCTTAAATGACAGATCTCTGCAAAATGAGTGAAGGCAAGGGAATTTCTTAC 1484
 DB 1293 AAACTCTATAGTCTTAAATGACAGATCTCTGCAAAATGAGTGAAGGCAAGGGAATTTCTTAC 1352
 QY 1485 AGTCCCTTGGTAAATTTAGGCACTCTTTTACGAAGTTCAGATATCTTAAGAAAGAT 1544
 DB 1353 AGTCCCTTGGTAAATTTAGGCACTCTTTTACGAAGTTCAGATATCTTAAGAAAGAT 1412
 QY 1545 TGAAGTATACAGATATGCTTGAATTTGATTCACCTCACAGTTCAGGAGATGTGACAT 1604
 DB 1413 TGAAGTATACAGATATGCTTGAATTTGATTCACCTCACAGTTCAGGAGATGTGACAT 1472
 QY 1605 TGGAAAAATGTTTCAATTAAGGGAACCGTTTATCATCTTCAAAATCATGTGTGACAGAT 1664
 DB 1473 TGGAAAAATGTTTCAATTAAGGGAACCGTTTATCATCTTCAAAATCATGTGTGACAGAT 1532
 QY 1665 TGATATCCACCTGGAGCAGTATTAGAGAACAGATTTGTGCTGGAACCTTCGCATCTT 1724
 DB 1533 TGATATCCACCTGGAGCAGTATTAGAGAACAGATTTGTGCTGGAACCTTCGCATCTT 1592
 QY 1725 GGACCACTGAAATGAAAAATACTGTGGAACCTTAAATAATGGGTAGTCTTCTTACAATGA 1784
 DB 1593 GGACCACTGAAATGAAAAATACTGTGGAACCTTAAATAATGGGTAGTCTTCTTACAATGA 1652
 QY 1785 AATGTTCTTAGGATTTTAAATAGGCAAGTACTTTTACTATGTTACTGACCTTCAGTG 1844
 DB 1653 AATGTTCTTAGGATTTTAAAGGTAAGGCACTTTTACTATGTTACTGACCTTCAGTG 1712
 QY 1845 TTGATTTTAAATAGAGTCTTCTGCACTATGCTTTAGTCTTAAGAAAGCAGACAGATGGA 1904
 DB 1713 TTGATTTTAAATAGAGTCTTCTGCACTATGCTTTAGTCTTAAGAAAGCAGACAGATGTT 1772
 QY 1905 GCAATACCTTCTCTTCTTTGAAGAGAAATCCCAAAAGTCTTCTTCTTAA 1953
 DB 1773 GCAATACCTTCTCTTCTTTGAAGAGAAATCCCAAAAGTCTTCTTCTTAA 1819

ID ABL62448 standard; DNA; 1823 BP.
AC ABL62448;
XX
XX
DT 15-MAY-2002 (first entry)
DE Colon adenocarcinoma related gene sequence SEQ ID NO:785.
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233677P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234034P.
XX 20-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234924P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 25-SEP-2000; 2000US-235280P.
XX 25-SEP-2000; 2000US-235637P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.
XX 28-SEP-2000; 2000US-236028P.
XX 28-SEP-2000; 2000US-236032P.
XX 28-SEP-2000; 2000US-236033P.
XX 28-SEP-2000; 2000US-236034P.
XX 28-SEP-2000; 2000US-236109P.
XX 28-SEP-2000; 2000US-236111P.
XX 29-SEP-2000; 2000US-236842P.
XX 29-SEP-2000; 2000US-236891P.
XX 02-OCT-2000; 2000US-237172P.
XX 02-OCT-2000; 2000US-237173P.
XX 02-OCT-2000; 2000US-237278P.
XX 02-OCT-2000; 2000US-237294P.
XX 02-OCT-2000; 2000US-237295P.
XX 02-OCT-2000; 2000US-237316P.
XX 03-OCT-2000; 2000US-237425P.
XX 03-OCT-2000; 2000US-237598P.
XX 03-OCT-2000; 2000US-237604P.
XX 03-OCT-2000; 2000US-237606P.
XX 03-OCT-2000; 2000US-237608P.
XX 01-NOV-2000; 2000US-244867P.
XX 01-NOV-2000; 2000US-245084P.
XX
XX (AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
PT

PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 785; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX oesophageal, ovarian, kidney, prostate or pancreatic cancer,
XX adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX Sequence 1823 BP; 593 A; 337 C; 402 G; 491 T; 0 other;
XX
XX Query Match 78.1%; Score 1676.6; DB 24; Length 1823;
XX Best Local Similarity 98.5%; Pred. No. 0;
XX Matches 1703; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
XX
XX 225 AGATCTTAGCAAGCAATGCTCTCAAGATGGTCTCTCAGTTCAGTTCAGAGAGTATTCGGCA 284
XX
XX 93 AGATCTTAGCAAGCAATGCTCTCAAGATGGTCTCTCAGTTCAGTTCAGAGAGTATTCGGCA 152
XX
XX 285 AGAGCTTAGAATATCTGTGAAGAGCACTAGAAAAAATCTCACCACAGCATCATCACA 344
XX
XX 153 AGAGCTTAGAATATCTGTGAAGAGCACTAGAAAAAATCTCACCACAGCATCATCACA 212
XX
XX 345 TGAATTTGAGCACACCAAAAAAGACCTCGATGGATTTTCGGAAGCTATTTTCATAGATTTT 404
XX
XX 213 TGAATTTGAGCACACCAAAAAAGACCTCGATGGATTTTCGGAAGCTATTTTCATAGATTTT 272
XX
XX 405 GCAAGAAAAAGGGGCTTCTGTGGATTGGGAAAAATCCAGAGACCCCTGGAAGATTCGAT 464
XX
XX 273 GCAAGAAAAAGGGGCTTCTGTGGATTGGGAAAAATCCAGAGACCCCTGGAAGATTCGAT 332
XX
XX 465 TCAACCCCTATGAAGAGATAAAGCCAGGGGCTTCTGTATATATATCTTCGGTGTGA 524
XX
XX 333 TCAACCCCTATGAAGAGATAAAGCCAGGGGCTTCTGTATATATATCTTCGGTGTGA 392
XX
XX 525 CAAACTAGTGGTGAACTCAATGGTGGTTTGGGAACCCAGCATGGGCTGCAAGGCC 584
XX
XX 393 CAAACTAGTGGTGAACTCAATGGTGGTTTGGGAACCCAGCATGGGCTGCAAGGCC 452
XX
XX 585 TAAAGTCTGATTTGGTGGAGAAATGAGAAATACCTTTCTGGATCTGACTGTTTCAGCAAT 644
XX
XX 453 TAAAGTCTGATTTGGTGGAGAAATGAGAAATACCTTTCTGGATCTGACTGTTTCAGCAAT 512
XX
XX 645 TGAACATTTGATAAACCCTACAATACAGATGTTCTCTTCTTTTAAAGAACTCTTTAA 704
XX
XX 513 TGAACATTTGATAAACCCTACAATACAGATGTTCTCTTCTTTTAAAGAACTCTTTAA 572
XX
XX 705 CACGGATGAAGATACCAAAAAAATACTACAGAAATGAGAAATCTTCGTTGTAATCTTA 764
XX
XX 573 CACGGATGAAGATACCAAAAAAATACTACAGAAATGAGAAATCTTCGTTGTAATCTTA 632
XX
XX 765 CACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTACTTCTTGTAGCAA 824
XX
XX 633 CACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTACTCGGCTGTAGCAA 692
XX
XX 825 GGACGTGTTCTACTCAGGGGAAAAATACAGAGCTTTGGTACCCTCCAGCTCATGGTAT 884
XX
XX 693 GGACGTGTTCTACTCAGGGGAAAAATACAGAGCTTTGGTACCCTCCAGCTCATGGTAT 752
XX
XX 885 TTACGCCAGTTTCTCAACTCTGGATTCGTTGATACCTTTATAGGAGGCAAGAGTA 944

Db 753 TTACGCCAGCTTCTACAACTCTGGATTGCTTGATACCTTTATAGGAGGACGCAAGATGTA 812
QY 945 TATTTTGTGTCTAACATAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCA 1004
Db 813 TATTTTGTGTCTAACATAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCA 872
QY 1005 TCTATGTAACCCACCACCAATGGAAGCGTGTGAATTTGTCATGGAGTCAACAAATAAACA 1064
Db 873 TCTAATCAACCCACCACCAATGGAAGCGTGTGAATTTGTCATGGAGTCAACAAATAAACA 932
QY 1065 ACCTGCAGATGTAAAGGGCGGACACTCACTCAATATGAAGCAAACTGAGACTGGTGA 1124
Db 933 ACCTGCAGATGTAAAGGGCGGACACTCACTCAATATGAAGCAAACTGAGACTGGTGA 992
QY 1125 AATTTGCTCAAGTGCACCAAGCACTAGTAGAGTTCAGTCTGTATCAAAAGTTCAAAAT 1184
Db 993 AATTTGCTCAAGTGCACCAAGCACTAGTAGAGTTCAGTCTGTATCAAAAGTTCAAAAT 1052
QY 1185 ATTTAATACAAACCACTATGATTTCTCTTGCAGAGTTAAAGACTGCAGGACAAA 1244
Db 1053 ATTTAATACAAACCACTATGATTTCTCTTGCAGAGTTAAAGACTGCAGGACAAA 1112
QY 1245 TGCCATTGACATGGAATCATTTGTAATGCAAGACTTTTGGATGGAGGCTCGAATGTCAT 1304
Db 1113 TGCCATTGACATGGAATCATTTGTAATGCAAGACTTTTGGATGGAGGCTCGAATGTCAT 1172
QY 1305 TCAATTAGAACTGCAAGTGGGGCTGCCATCAAAAGTTTGAATTTCTTAGGTATTTAA 1364
Db 1173 TCAATTAGAACTGCAAGTGGGGCTGCCATCAAAAGTTTGAATTTCTTAGGTATTTAA 1232
QY 1365 TGTGCCAAGGAGCGTTTCTGCTGTCAAAACCACTCAGATCTCTTCTGGTGTATGTC 1424
Db 1233 TGTGCCAAGGAGCGTTTCTGCTGTCAAAACCACTCAGATCTCTTCTGGTGTATGTC 1292
QY 1425 AAACCTCTATAGTCTTAATGCAAGATCTCTGCAATGAGTGAAGGCGGAATTTCCCTAC 1484
Db 1293 AAACCTCTATAGTCTTAATGCAAGATCTCTGCAATGAGTGAAGGCGGAATTTCCCTAC 1352
QY 1485 AGTGCCCTTTGGTTAAATAGGAGGTTCTTTTACGAAGTTCAAGATTTCTAAGAAGATT 1544
Db 1353 AGTGCCCTTTGGTTAAATAGGAGGTTCTTTTACGAAGTTCAAGATTTCTAAGAAGATT 1412
QY 1545 TGAAGATATACAGATATCTTGAATTTGATCACTCAGTTTCAGGAGATGTGACATT 1604
Db 1413 TGAAGATATACAGATATCTTGAATTTGATCACTCAGTTTCAGGAGATGTGACATT 1472
QY 1605 TGAAGATATCTTCAATTAAGGGAACGGTTATCATCTTGCATAATCATGGTGACAGATT 1664
Db 1473 TGAAGATATCTTCAATTAAGGGAACGGTTATCATCTTGCATAATCATGGTGACAGATT 1532
QY 1665 TGATATCCACCTGGAGAGATATTAGAGAACAGATTGTCTGGAACCTTCGCAATCTT 1724
Db 1533 TGATATCCACCTGGAGAGATATTAGAGAACAGATTGTCTGGAACCTTCGCAATCTT 1592
QY 1725 GGACCACTGAATGAATAATCTGTGACACTTAAATATATGGCTAGTTTCTTACAAATGA 1784
Db 1593 GGACCACTGAATGAATAATCTGTGACACTTAAATATATGGCTAGTTTCTTACAAATGA 1652
QY 1785 AATGTTCTCTAGGATTTCAAAATAGGACGATCTTTTACTATGTTACTGTACCTCGAGTG 1844
Db 1653 AATGTTCTCTAGGATTTAGGCACATAAAGGTACTTTACTATGTTACTGTACCTCGAGTG 1712
QY 1845 TTGATTTTAAATAGAGTTTCTGCAGTATGCTTTTATGCTTAAGAAAGCAAGATGA 1904
Db 1713 TTGATTTTAAATAGAGTTTCTGCAGTATGCTTTTATGCTTAAGAAAGCAAGATGA 1772
QY 1905 GCAATACTTCTCTTTGAGAGNATCCCAAGTTAGTTCATCTTAA 1953
Db 1773 GCANTACTTCTCTTTGAGAGAGA--TCCCAAGTTAGTTCATCTTAA 1819

AAC10871

ID AAC10871 standard; cDNA; 853 BP.

XX AAC10871;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 14946.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 14946; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 853 BP; 261 A; 160 C; 232 G; 200 T; 0 other;

Query Match. 38.6%; Score 828; DB 21; Length 853;

Best Local Similarity 99.9%; Pred. NO. 4e-215;

Matches 839; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GTGTTTAGAGAAAGTAGGGCTGTGGGTGTGGGAGCGGCTGACGGGTGACAAGGG 60

Db 15 GTGTTTAGAGAAAGTAGGGCTGTGGGTGTGGGAGCGGCTGACGGGTGACAAGGG 74

QY 61 GGGTTTAGCAGCTGGGCTCGACCGTTAGGAGGGGCTCAAGGTGTGCATGTGTAGGAA 120

Db 75 GGGTTTAGCAGCTGGGCTCGACCGTTAGGAGGGGCTCAAGGTGTGCATGTGTAGGAA 134

QY 121' GAGAGAGAGAGAGAGGGGCTCGAGGTGACATTTTACGCTGCGAGCTTCTCCGG 180

Db 135 GAGAGAGAGAGAGAGGGGCTCGAGGTGACATTTTACGCTGCGAGCTTCTCCGG 194

QY 181 GCGCCATAAAGCGCCCAATTTCCAGCTGTCTAAAGAGAGAGAGATCTTAGCAAGCA 240

Db 195 GCGCCATAAAGCGCCCAATTTCCAGCTGTCTAAAGAGAGAGAGATCTTAGCAAGCA 254

QY 241 ATGCTCTCAAGATGGTCTTCTCAGTTCCAAGAGATCATTCGGCAAGAGCTAGAATTATCT 300

Db 255 ATGTCTCAAGATGGTCTTCTCAGTTCAGTCCAGAGATCAATTCGGCAGAGCTAGAAATATCT 314
 QY 301 GTGAAGAAGAACTAGAAAAAATCTACACACAGCATCATCATGAATTTGAGCACACC 360
 Db 315 GTGAAGAAGAACTAGAAAAAATCTACACACAGCATCATCATGAATTTGAGCACACC 374
 QY 361 AAAAAGACCTGGATGGATTTCCGAAGCTATTTCATAGATTTTTCAGAAAAAGGGGCT 420
 Db 375 AAAAAGACCTGGATGGATTTCCGAAGCTATTTCATAGATTTTTCAGAAAAAGGGGCT 434
 QY 421 TCTGTGGATTTGGGAAAAATCCAGAGACCCCTCGAAGATTCCGATTCACCCCTATGAAAG 480
 Db 435 TCTGTGGATTTGGGAAAAATCCAGAGACCCCTCGAAGATTCCGATTCACCCCTATGAAAG 494
 QY 481 ATAAAGCCAGGGGCTTGGCTGATAATATATCTTCCGTGTGAAACAACTAGTGTGGTG 540
 Db 495 ATAAAGCCAGGGGCTTGGCTGATAATATATCTTCCGTGTGAAACAACTAGTGTGGTG 554
 QY 541 AAACCTCAATGGTGGTGGGAACCCAGCATGGGCTGAAAAGGCCCTAAAAGTCTGATTGGT 600
 Db 555 AAACCTCAATGGTGGTGGG-ACCAGCATGGGCTGAAAAGGCCCTAAAAGTCTGATTGGT 613
 QY 601 GTGAGGAATGAGAATACCTTTCTGGATCTGACTGTTTCAGCAAAATTCGAATTTGAATAA 660
 Db 614 GTGAGGAATGAGAATACCTTTCTGGATCTGACTGTTTCAGCAAAATTCGAATTTGAATAA 673
 QY 661 ACCTACAATACAGATGTTCTCTTGTGTTTAAATGAACCTTTTAAACCGGATGAAGATACC 720
 Db 674 ACCTACAATACAGATGTTCTCTTGTGTTTAAATGAACCTTTTAAACCGGATGAAGATACC 733
 QY 721 AAAAAAATCTACAGAAGTACAAATCAATGTCGTGTGAAAAATCTACATTTCAATCAAGC 780
 Db 734 AAAAAAATCTACAGAAGTACAAATCAATGTCGTGTGAAAAATCTACATTTCAATCAAGC 793
 QY 781 AGTACCCGAGGATTAATAAGATCTTTACTTCTGTCAGCAAGACGCTGCTTACTCA 840
 Db 794 AGTACCCGAGGATTAATAAGATCTTTACTTCTGTCAGCAAGACGCTGCTTACTCA 853

RESULT 4
 AAI97888/c
 ID AAI97888 standard; cDNA; 709 BP.
 XX
 AC AAI97888;
 XX
 DT 13-NOV-2001 (first entry)
 XX
 DE Human neuroblastoma expressed polynucleotide SEQ ID NO 3963.
 XX
 KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200166719-A1.
 XX
 PD 13-SEP-2001.
 XX
 PE 02-MAR-2001; 2001WO-JP01629.
 XX
 PR 07-MAR-2000; 2000JP-0159195.
 XX
 PA (CHIB-) CHIBA PREFECTURE.
 PA (HISM) HISAMITSU PHARM CO LTD.
 XX
 PI Nakagawara A;
 XX
 DR WPI; 2001-565584/63.
 XX
 PT Nucleic acids originating in gene expressed in human neuroblastoma,
 PT useful as probe or primer in diagnosing prognosis of human
 PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
 PT for anti-cancer agents
 XX

PS Claim 1; Page 2914-2915; 2979pp; Japanese.
 XX
 CC The invention relates to novel genes (AAI93926-AAI97963) expressed in
 CC human neuroblastoma. The nucleic acids are applicable as a probe or
 CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
 CC susceptibility indicators or tumour markers for anti-cancer agents. The
 CC gene information for diagnosing prognosis is related to factors similar
 CC to that for N-myc and TrkA genes.
 XX
 SQ Sequence 709 BP; 210 A; 136 C; 119 G; 239 T; 5 other;
 Query Match 31.4%; Score 675; DB 22; Length 709;
 Best Local Similarity 99.9%; Pred. No. 1.8e-173;
 Matches 675; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1394 AAACACATCAGATCTTCTGCTGGTGAATTCCTACAGTCCCTTGGTTAAATTTAGCAGTCTC 1453
 Db 707 AAACACATCAGATCTTCTGCTGGTGAATTCCTACAGTCCCTTGGTTAAATTTAGCAGTCTC 648
 QY 1454 TGCAATCAGTGAAGGCGGAATTCCTACAGTCCCTTGGTTAAATTTAGCAGTCTC 1513
 Db 647 TGCAATCAGTGAAGGCGGAATTCCTACAGTCCCTTGGTTAAATTTAGCAGTCTC 588
 QY 1514 TTACGAAGTTCAGATTTATCTAAGAAGATTCTAAGAAGATTCTAAGAAGATTCTAAGAAG 1573
 Db 587 TTACGAAGTTCAGATTTATCTAAGAAGATTCTAAGAAGATTCTAAGAAGATTCTAAGAAG 528
 QY 1574 ATCACTCAGTTCAGGAGATGACATTTGGAAAAAATGTTTCATTTAAAGGGAACCG 1633
 Db 527 ATCACTCAGTTCAGGAGATGACATTTGGAAAAAATGTTTCATTTAAAGGGAACCG 468
 QY 1634 TTATCATCATTTGCAATCATGTGACAGAAATGATATCCACCTGGAGCAGTATTAGAGA 1693
 Db 467 TTATCATCATTTGCAATCATGTGACAGAAATGATATCCACCTGGAGCAGTATTAGAGA 408
 QY 1694 ACAAGATTGTCGTGGAAACCTTCGCATCTTGGACCACTGAAATGAAAAAATCTGTGGAC 1753
 Db 407 ACAAGATTGTCGTGGAAACCTTCGCATCTTGGACCACTGAAATGAAAAAATCTGTGGAC 348
 QY 1754 ACTTAAATTAAGGGCTAGTTCCTTCAATGAATGTTCTCTAGGATTCCTAAATAGGCAG 1813
 Db 347 ACTTAAATTAAGGGCTAGTTCCTTCAATGAATGTTCTCTAGGATTCCTAAATAGGCAG 288
 QY 1814 GTACTTTTACTATGTTTACTGTACCTGTGAGTGTGATTTTAAATAGAGTTCCTGCACT 1873
 Db 287 GTACTTTTACTATGTTTACTGTACCTGTGAGTGTGATTTTAAATAGAGTTCCTGCACT 228
 QY 1874 ATGCTTTTACTATGTTTACTGTACCTGTGAGTGTGATTTTAAATAGAGTTCCTGCACT 1933
 Db 227 ATGCTTTTACTATGTTTACTGTACCTGTGAGTGTGATTTTAAATAGAGTTCCTGCACT 168
 QY 1934 CAAGAAGTTAGTTCATCTTAAAGTGAATATGTTTAAATCTTAAACTGGCACTTGA 1993
 Db 167 CAAGAAGTTAGTTCATCTTAAAGTGAATATGTTTAAATCTTAAACTGGCACTTGA 108
 QY 1994 AGAACTTTTAAAGAGGCTCAATGATGATCACTTTGAAATGCTTTGATTTCAAAAAATA 2053
 Db 107 AGAACTTTTAAAGAGGCTCAATGATGATCACTTTGAAATGCTTTGATTTCAAAAAATA 48
 QY 2054 AAGCAGTGAAGCAATA 2069
 Db 47 AAGCAGTGAAGCAANA 32
 RESULT 5
 AAI97887
 ID AAI97887 standard; cDNA; 744 BP.
 XX
 AC AAI97887;
 XX
 DT 13-NOV-2001 (first entry)
 XX
 DE Human neuroblastoma expressed polynucleotide SEQ ID NO 3962.

QY 1017 ACCCAATGGAAGAAAGCTGTGAATTTTGTCAATGAAGTCAAAAATAAACAACGTCGACATGT 1076
Db 967 GCGGGCCACTACTCCAGTGGAGTTTGTCAATGAGGTCCCGACAGACCCGTCGTGACGT 1026
QY 1077 AAAGGGCGGGACACTCACTCAATATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGT 1136
Db 1027 TAAGGGTGTACTCTCATTCATCAAAATGGAACAAAGCTGGCCCTGTTGGAAATTTGCTCAAGT 1086
QY 1137 GCCAAAAGCAATGTAGACGATTCAGTCTGTATCAAAAGTTCAAAATTTAAATACAAA 1196
Db 1087 GCCACGAGCATGTGGATGACTTTAAGTCGGTTAAGACCTTTCAAGTTCTTTTAAACACCAA 1146
QY 1197 CAACCTATGATTTCTTCTTGCAGCAGTTAAAGACTGACAGGAGCAAAATGCCATTGACAT 1256
Db 1147 CAACATTTGGGCAACCTGGCAGCCATTCATCGGTTTTCGTGAGCTACCTTAAACAT 1206
QY 1257 GGAATCATTTGTGAATGCAAGACTTTGGATGAGGCTTGAATGTCAATTAATTAAGAAC 1316
Db 1207 GGAATATTATTGTGAACAACAGACGCTGGAGAAGCAACCCGTGTCAATTCAGCTGGAGAC 1266
QY 1317 TCAGTAGGGCTGGCATCAAAAGTTTGAAGATTTCTTAGTATTAATGTCGCAAGGAG 1376
Db 1267 TGCCTGGGTGGGCGCATGAAGTGTCTTGTATGGCGCATCGGCATCAATGTGCGCGTTC 1326
QY 1377 CCGTTTCTGCTGTCAAAACACACATCAGATCTTGTGCTGATGTCAAACTCTATAG 1436
Db 1327 CCGTTTCTGCGGTGAAGAAATCTTCAGATCTGCTGCTGGTCAATGTCMAATTTATATAC 1386
QY 1437 TCTTAATGACGATCTCTGACAAATGAGTGAAGGAGGGAATTTCTTACAGTGCCTTGT 1496
Db 1387 GCTTAAGAACGGCAGCTGTGTAATGTGCGCCACAGCGAATGTTCCCTACACGCGCTGT 1446
QY 1497 TAAATTT---AGCAGTTCTTTTACGAAGTTTCAAGATTTATCAAGAAATTTTGAAGTAT 1553
Db 1447 CAAGCTGGGAGAGAACCACTTCTCGAAGTCAAGGAGTTCTCTGGGTGCTGTTGGCAACAT 1506
QY 1554 ACCAGATATGCTGAATTCGATCACCCTCACAGTTTCAGGAGATGTGACATTTGGAAAAA 1613
Db 1507 CCGGACATCATTTGAGCTGGATCATTGACCGTGAGCGGAGATGTGACTTTCGCTGGTG 1566
QY 1614 TGTTTCATTAAGGGAAACGGTTATCATATTCGCAATCATGTTGAGAGAAATTTGATATCC 1673
Db 1567 CGTTCTTTTGGCGGCACTGTGATCATCATTTGTAACCATGGCATCGTATCGATATCC 1626
QY 1674 ACTGAGCAGATTTAGAGAACAGATTTGTGCTGGAACCTTCGATCTTGGACCACTG 1733
Db 1627 AGCTGGAGCCATTTCTGGAGAACAGATTTGATCTGGCAATATGCGCATCTTGGACCACTA 1686
QY 1734 AATGAA 1740
Db 1687 AATCAAA 1693

RESULT 8
ABL07483
ID ABL07483 standard; cDNA; 2222 BP.
XX AC ABL07483;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 16931.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-655860/75.
DR P-PSDB; ABB63380.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 16931; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2222 BP; 584 A; 494 C; 529 G; 615 T; 0 other;
Query Match 25.0%; Score 535.8; DB 23; Length 2222;
Best Local Similarity 63.7%; Pred. No. 2.6e-135;
Matches 883; Conservative 0; Mismatches 492; Indels 12; Gaps 4;
QY 360 CAAAAAACACCTCGATGATTTCCGAAGCTATTTTCATAGATTTTTCGAAGAAAAGGGGCC 419
Db 348 CAAGCAGAAATGGTCTGTTTCGCCGATCTGTCGACGTTTCATTTCAGGAAGAGGGTCC 407
QY 420 TTCTGTGATTTGGGAAAAAATCCAGACCCCTCGAAGATTCGATTCAACCCCTATGAATA 479
Db 408 CGCATTTGCACTGGAACAAGATTCGAAGCTGCTCGAAGATCTGTGATGAACCTACTCGAA 467
QY 480 GATAAGCCAGGGGCTTCCCTGATAATATATCTTCCGTGTTGAACAACTAGTGGTGT 539
Db 468 TCTTAAGTC---GCCAAGAACGACGATCGGCAACATTTGGATAAATTCGTGTGTCAT 524
QY 540 GAAACTCAATGGTGGTTTGGGAACCCAGCATGGGCTGCAAGGCCCTAAAGTCTGATTGG 599
Db 525 CAAGCTGAACGGTGGTTTGGGCACCTCAATGGTTGCCATGGTCCCAAGAGTGTGATTCC 584
QY 600 TGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTTCAGCAATTTGAACATTTGAATAA 659
Db 585 TGTGCTTCGGATCTTACTTTCTGGATCTGACCGTGCAGCAAAATCGAGCATCTCAACAA 644
QY 660 AACCTACAAATACAGATGTTCTCTTTGTTTAAATGAATCTTTTAAACGAGTGAAGATAC 719
Db 645 GACCTACGATGCAATGTTCCGCTGTACTGATGAATCTCTTTAAACGAGGAGATAC 704
QY 720 CAAAAAATACTACAGAAAGTACAATCATTTGTGCTGTGAAATCTTACACTTTTCAATCAAG 779
Db 705 CGAAGATCGTGGCCNAAGTACAAGGATTCGCTGTGAGATCCACACCTTCAACAGAG 764
QY 780 CAGGTACCGAGGATTAATAAGAAATCTTTACTTCTGTAGCAAGGAGCTGTCTACTC 839
Db 765 CTGCTTCCCGCATTTAGTCGCGACACTTACCTGCGGTGGCCAAAGACTTTTGTATGTCGA 824
QY 840 AGGGGAATATACAGAAAGCTTGGTACCTCCAGTCTGATGATATTTAGCCAGTTTCTA 899
Db 825 AAAGG---ATATGGAAGCTGTGTATCCACTGGACACCGTGTACTTCTAGCATACCTTCG 881
QY 900 CAACTCTGATTTGATACCTTTATAGGAGAGGCAAGAGTATATTTTGTGTCTAA 959

Db 882 CAACCTCTGGCTGCTGAAGAGATTTCATCGAGGAGGCGCGAGTATTGTTCTTGTGTCAC 941
Qy 960 CATAGATAATCTGGGTGCGACAGTGGATCTGTATATCTTAATCATCTAAATG---AACCC 1016
Db 942 CATGATTAACCTGGGCGCACCGTGCATCTGAACATCTCAATAAGCTGGTGGCGGAGGA 1001
Qy 1017 ACCCAATGGAAACGCTGTGAATTTGTTCATGGAAGTCAACAATAAACAACGTCGAGATGT 1076
Db 1002 CGGGGCCACTACTCCAGTGGAGTTTGTTCATGGAGGTCAACGACAGACCCGTCGTGACGT 1061
Qy 1077 AAAGGGCGGACACTCACTCAATTAAGAGCAAACTGAGACTGGTGGAAATTTGCTCAAGT 1136
Db 1062 TAAGGTGGTACTCTTCATTAAGTGAACAAAGCTGCGCTGTTGGAAATTTGCTCAAGT 1121
Qy 1137 GCCAAAGACACATGTAGACGATTCAGTCTGTATCAAAAGTTCAAAATTTAATAACAAA 1196
Db 1122 GCCACGAGCATGTGGATGACTTTAAGTCGGTTAAGACCTTCAAGTTCTTTTAACACCA 1181
Qy 1197 CAACCTATGGATTTCTTGTGAGCAGTTAAAGACTGCGAGGAGCAAAATGCCAATTGACAT 1256
Db 1182 CAACATTTGGGCGCAACCTGGAGCCATTGATCGCGTGTGGGTGAGCGTACCTTAAACAT 1241
Qy 1257 GGAATCATTTGTAATGCAAGACTTTTGGATGGAGCGCTGATGTCATTCATTAATTAGAAC 1316
Db 1242 GGAATATTGTGAACAAACAGACCGCTGGAGAACCGCGTGTCACTTACGTGGAGAC 1301
Qy 1317 TGCAGTAGGGGTGCGCATCAAAAGTTTGTGAGAAATTTCTAGTATTAATGTGCCAAGGAG 1376
Db 1302 TGCTGTGGTGGCGCATGAGTGTGATGGCGCCATCGGCATCAATGTGCGCGTTC 1361
Qy 1377 CGTTTTCTGCTGTCAAAACACATCAGATCTTGTGTGGTGTGATGTCGCAACCTCTATAG 1436
Db 1362 CGTTTTCTGCGGTGAAGAAATCTTCAGATCTGCTGTGGTGTGATGTCGCAAAATTTATATAC 1421
Qy 1437 TCATTAATCGAGATCTTCGACATGAGTCAAAAGCGGAAATTTCTACAGTGCCTTGGT 1496
Db 1422 GCTTAAGAACCGCAGCTGGTGTAAATGTCGCGACAGCGAAATGTTCCCTACACGCGCTCGT 1481
Qy 1497 TAAATTT---AGGCAGTCTTTTACGAAGTTTCAAGATTATCTAAGAAGATTGAAAGTAT 1553
Db 1482 CAAGCTGGGAGAACCACTTCTCAAGGTCAAGGATGCTTGGTGGCTTGGCAACAT 1541
Qy 1554 ACCAGATATGTTGAATTTGATCCTCAGTTCACAGTTCAGAGATGTGACATTTGGAAAAA 1613
Db 1542 CCGGACATCATTTGAGCTGGATCACTTGACCGTGAGCGGAGATGTGACTTTCGGTGG 1601
Qy 1614 TGTTTCATTAAGGGAACGGTTATCATCATTCGAAATCATGTCGACAGAAATGATCCC 1673
Db 1602 CGTTTTCTTTCGCGGCACTGTGATCATCATTCGTAACCACTGCGATCGATATCCC 1661
Qy 1674 ACCTGGAGCAGTATTAGAGAACAGATTGTGTCTGGAACCTTCGCACTCTTGGACCACTG 1733
Db 1662 AGCTGGGCCATTCTGGAGAACAGATTGTATCTGGCAATATGCGCACTTGGACCCTA 1721
Qy 1734 AATGAA 1740
Db 1722 AATCAAA 1728

RESULT 9
ABQ58840
ID ABQ58840 standard; cDNA; 701 BP.
XX
AC ABQ58840;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:2535.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.

XX WO200229086-A2.
PN 11-APR-2002.
XX 02-OCT-2001; 2001WO-US30732.
PF 02-OCT-2000; 2000US-237271P.
PR (FARB) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thagalingam A, Lewis ME;
XX WPI; 2002-426115/45.
DR
XX New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
SQ Sequence 701 BP; 214 A; 129 C; 142 G; 193 T; 23 other;

Query Match 24.2%; Score 519.2; DB 24; Length 701;
Best Local Similarity 94.4%; Pred. No. 4.8e-131;
Matches 575; Conservative 0; Mismatches 29; Indels 5; Gaps 4;
Qy 860 GGTACCCCTCCAGGTGATGTTTACGCCAGTTTCTACAACTCTGGATTGCTTGATA 919
Db 4 GGTACCCCTCCAGGTGATGTTTACGCCAGTTTCTACAACTCTGGATTGCTTGATA 63
Qy 920 CTTTATAGGAGGCAAGAGATATATTTTGTGCTTAACATAGATATCTGGGTGCCA 979
Db 64 CTTTATAGGAGGCAAGAGATATATTTTGTGCTTAACATAGATATCTGGGTGCCA 123
Qy 980 CAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCCACTGAAACGCTGTGAT 1039
Db 124 CAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCCACTGAAACGCTGTGAT 183
Qy 1040 TTGTCATGGAAGTCAAAATAAACAACGTCGAGATGTAAGCGGGGACACTCACTCAAT 1099
Db 184 TTGTCATGGAAGTCAAAATAAACAACGTCGAGATGTAAGCGGGGACACTCACTCAAT 243
Qy 1100 ATGAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCAAAACACATGTAGACGAT 1159
Db 244 ATGAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCAAAACACATGTAGACGAT 303
Qy 1160 TCAAGTCTGTATCAAAAGTTCAAAATTTAATACAAACCTATGGATTTCTTTCGAC 1219
Db 304 TCAAGTCTGTATCAAAAGTTCAAAATTTAATACAAACCTATGGATTTCTTTCGAC 363
Qy 1220 C-AGTTAAAGACTGCGAGCAAAATGCAATTTGATGGAATCATTTGTAATGCAAAAG 1278
Db 364 CAAGTTAAAGACTGCGAGCAAAATGCAATTTGATGGAATCATTTGTAATGCAAAAG 423

QY 1279 ACTTTGGATGGAGGCTGAATGCTCAATTAATAGAACTGCAGTAGGGCTGCCATCAA 1338
 DB |||||
 424 ACTTTGGATGGA-NCTTGATGCTCAATTAATAGAACTGCAGTAGGGCTGNCATC-AA 481
 QY 1339 ACTTTGAGAACTCTAGGTAATTAATGTCGAAGAGCGTTTCTGCTGTCAAAACC 1398
 DB |||||
 482 AGTTTGGAAATCTCTAAGNATTAATGTCGAAGAGCGCTTCTGCTGTCAAAACA 541
 QY 1399 ACATCAGATCTCTGCTGATGTCGAACCTCTATAGTCTTAATGCGAGGATCTCTGACA 1458
 DB |||||
 542 CNTT--AGACTCTTGCTNGATGTCGAACCTCTATAGTCTTAATGCGAGGATCTCTGACA 599
 QY 1459 ATGAGTGAA 1467
 DB 600 TGATGNAAA 608

RESULT 10

ABK62970/c
 ID ABK62970 standard; cDNA; 671 BP.

XX AC ABK62970;

XX DT 18-JUN-2002 (first entry)

XX DE Rat sequence differentially expressed in response to a hepatotoxin #877.

XX KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 differential expression; centrilobular necrosis; steatosis.

XX OS Rattus norvegicus.

XX PN -WO200210453-A2.

XX PD 07-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23872.

XX PR 31-JUL-2000; 2000US-222040P.

XX PR 02-NOV-2000; 2000US-244880P.

XX PR 11-MAY-2001; 2001US-290029P.

XX PR 15-MAY-2001; 2001US-290645P.

XX PR 22-MAY-2001; 2001US-292336P.

XX PR 06-JUN-2001; 2001US-295798P.

XX PR 13-JUN-2001; 2001US-297457P.

XX PR 19-JUN-2001; 2001US-298884P.

XX PR 09-JUL-2001; 2001US-303459P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Mendrick D, Porter MW, Johnson KP, Castle AL, Elashoff MR;

XX DR WPI; 2002-241625/29.

XX PT Predicting toxic effects of compounds or the progression of these toxic
 effects by determining the changes in gene expression in tissues or
 cells exposed to the toxin and comparing these to gene expression in
 unexposed tissues or cells -

XX PS Claim 1; Seq ID No 877; 239pp; English.

XX CC The invention relates to methods for predicting toxic effects of
 compounds or the progression of these toxic effects by determining the
 global changes in gene expression in tissues or cells exposed to the
 toxin and comparing these to gene expression in unexposed tissues or
 cells. Also included are methods of predicting at least one toxic
 effect of a compound or progression of a toxic effect, preferably the
 hepatotoxicity of a compound, comprising detecting the level of
 expression in a tissue or cell sample exposed to the compound of two or
 more genes listed in the specification, where differential expression of
 the genes is indicative of at least one toxic effect or progression.
 The method can also be used to identify an agent which modulates the
 toxic response and predict cellular pathways that a compound modulates

CC in a cell. The methods utilise a set of at least two probes (on a solid
 support in kit form), where each of the probes comprises a sequence that
 specifically hybridises to a gene listed in the specification, a computer
 system comprising a database containing information identifying the
 expression level in a tissue or cell sample exposed to a hepatotoxin of a
 set of genes comprising at least two genes listed in the specification,
 and a user interface to view the information used to present information,
 CC identifying the expression level in a tissue or cell of at least one gene
 listed in the specification. The method is useful for elucidating global
 changes in gene expression and for identifying toxicity markers in
 tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 gene expression information may be used as diagnostic markers for the
 prediction or identification of the physiological state of tissue or cell
 sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.
 XX
 SQ Sequence 671 BP; 208 A; 127 C; 124 G; 210 T; 2 other;

Query Match 19.6%; Score 419.8; DB 24; Length 671;

Best Local Similarity 81.3%; Pred. No. 5.7e-104;

Matches 536; Conservative 0; Mismatches 114; Indels 9; Gaps 4;

QY 1403 CAGATCTCTTGTGCTGATGTCAAACTCTATAGTCTTAATGCGAGGATCTCTGACAATGA 1462

DB 671 CAGATCTCTTACTTGTGATGTCAAACTTACAGCTTACGCGAGATCTTGACCATGA 612

QY 1463 GTGAAACGGGAATTCCTACAGTGCCTTGGTTAAATAGGAGGATCTTTTACGAAGG 1522

DB 611 GTGAAACGGTGAATTCNTACAGTACCTTTGGTTAAATAGGAGGATCTTTTACCAAGG 552

QY 1523 TTCAGATTATCTAAGNAGATTGAAAGTATACACAGATATGCTTCAATTTGGATCACTCA 1582

DB 551 TTCAGATTATCTAAGNAGATTGAAAGTATACCCGATATGTTGAACTGGACCACTCA 492

QY 1583 CAGTTTCAGGAGATGTGACATTTGGAAAAAATGTTTCAATTAAGGGAACGGTTATCATCA 1642

DB 491 CGGTTTCAGGAGATGTGACATTTGGAAAAAATGTTTCAATTAAGGGAACAGTTATCATCA 432

QY 1643 TTGCAAAATCATGGTGACAGAAATGATATCCACCTGGAGGAGCAGTATTAGAGAACAGATTG 1702

DB 431 TTGCAAAATCATGGTGACAGAAATGATATCCACCTGGAGGAGCAGTATTAGAGAACAGATTG 372

QY 1703 TGCTTGGAAACCTTCGCATCTTGGACCACTGAAATGAAATAATCTGTGGACACTT---- 1757

DB 371 TATCTGGGAACCTTCGGATCTCGGACCACTGAAATGAGCACTGTGAGCTACACTTTCTTAC 312

QY 1758 AAATAATGGGCTAGTTTCTTACAATGAATGTTCTTAGGATTTCTAAATAGGCAGGTAC 1817

DB 311 TAATTTATGGGCTAAGAGTTTACAAATAAATGTCCTCAGATTATTAAGCAAGCAGGTAC 252

QY 1818 TTTA-CTATGTTACTACCCCTGCAGTGTGATTTTAAATAGAGTATTTCTGCGAGTATG 1876

DB 251 TTTATTTACTATGTTGACTCTGCACATATGATTTTAAAGTA--GTTTTCTGCAATGAG 194

QY 1877 CTTTATTAGTCTAGAAAGCAGACAGATGGAGCAATATCTTCTCTTTTGAAGAGATCCCAA 1936

DB 193 CTTTATTAGTCTAAGGAAAGCCTTAAACAAAGCAATATCTTCTCTTTGGATGAAAGCCTAG 134

QY 1937 AAGTTAGTCTCATCTTTAAAGTGCAATATTTGTTTAAATCTTTAAACTGGGCAACTTTGGAAGA 1996

DB 133 AAATTTAGGTCCTCTTAAGAGTGAATATTTATTTAATCTCAGATCGGGCCAGCTCTGTAGA 74

QY 1997 ACTTTTAAACAGAGGCTCAATGATGATCATCTTTGAATTCCTGTTGATTTCAAAAATAAA 2055

DB 73 TC-TTCAACAGAGGCTTCAGTGACTGCTCCCTTTGAATTTGATTTTCAAAAATAAA 16

RESULT 11

AAA67125

ID AAA67125 standard; DNA; 2045 BP.

XX AAA67125;
AC 31-OCT-2000 (first entry)
DT Pinus radiata UGP nucleotide sequence SEQ ID NO:80.
XX
XX Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KW plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
KW transgenic plant; ds.
XX
XX Pinus radiata.
OS
XX WO200022092-A2.
PN
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-NZ00169.
XX
XX 13-OCT-1998; 98US-0170862.
PR 11-AUG-1999; 99US-0148426.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Blokeberg LN;
PI
XX
XX WPI; 2000-339328/29.
DR P-PSDB; AAB16318.
XX
XX
XX New genes encoding proteins involved in a plant polysaccharide
PT biosynthetic pathway, useful for modulating or altering the
PT polysaccharide content, composition or structure of the plant
XX
XX Claim 1; Page 76; 301pp; English.
XX
XX The present invention describes isolated polynucleotides (PN) comprising
CC a sequence selected from one of 835 nucleotide sequences given in
CC AAA67073 to AAA67907, their (reverse) complements, sequences producing
CC an Expectation (E) value of 0.01 or less compared to the 835 sequences,
CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC 835 sequences or sequences that are degenerately equivalent or allelic
CC to the 835 sequences. The polynucleotides are used to modify the
CC activity of a polypeptide involved in a polysaccharide biosynthetic
CC pathway in the plant. They are especially used to modulate or alter the
CC polysaccharide content, composition or structure of the plant. AAB16268
CC to AAB16340 are proteins encoded by some of the polynucleotide sequence
CC given in the present invention.
XX
XX Sequence 2045 BP; 607 A; 387 C; 493 G; 558 T; 0 other;
SQ
Query Match 18.3%; Score 392.2; DB 21; Length 2045;
Best Local Similarity 57.9%; Pred. No. 3.2e-96;
Matches 785; Conservative 0; Mismatches 543; Indels 27; Gaps 4;
QY 353 AGCACACCAAAAAGACCTGGATGGAATTCGGAAGCTATTTTCATAGATTTTTCGAAGAA 412
DB 404 AGATCAGTGATAATGAGAAGGAAGGTTTGTGCGACTGTTTCTCGCTACCTCAGCGCG 463
QY 413 AGGGGCTTCTGTGATTTGGGGAATAATCCAGAGACCCCTCGAAGATTCGATTCACCCCT 472
DB 464 AAGAAGAGAAGATTGAATGGGGAATAATCAAGACCCCGACTGACGAGATCGTGGTTCCCT 523
QY 473 ATGAAAGATAAAGGCCAGGGG---CTTGCCTGATAATATATCTTCCGTTGTTGAACAAAC 529
DB 524 AGCATACGTTGGCAGCTTTGGGAGAAGATCCATCGGAACCAAGGAACTCTTGGCAAGC 583
QY 530 TAGTGGTGTGAACCTCAATGGTGGTTTGGGAACCCAGCATGGGCTGCAAGGCCCTTAAAA 589
DB 584 TTGTTGTGTTAAAGCTTAATGGCGTTTGGGAACCAACCATGGTTGTACTGGGCCCAAT 643
QY 590 GTCTGATTTGGTGAGGAATGAGATACCTTTCTGGATCTGATCTGTTCAGCAAAATGAAC 649
DB

DB 644 CCGTCATTGAAGTACGAAATGCCCTAACATTTCTGGACTTGATTTGTGAAGCAAAATAGACT 703
QY 650 ATTTGAATAAAACCTTACAATACAGATGTTCTCTCTTTTAAATGAATCTTTTAAACACGG 709
DB 704 CTCTGAAACAACAAATACGATTCGAAGTTCCGTTGGTCTGATGAATTCATTTAATACAC 763
QY 710 ATGAAGATACCAAAAAATATCTACAGAGTACAATCATTTGTCGTGTGAAAAATCTACATTT 769
DB 764 ATGATGATACAATTAAGATTGTAGAAAAAGTACTCTGGCTCAAAATATTGACATCCCATCT 823
QY 770 TCAATCAAAGCAGGTACCGGAGGATTAAATGAAGATCTTTTACTTCTCTGTAGCAAGGACG 829
DB 824 TCAACGAGTCAATATCCAGCATGGTGGCAGAAGATTTTACACCATGCGCCAA----- 877
QY 830 TGTCTTACTCAGGGGAAAAATACAGAAAGCTTGTGTACCTCCAGCTCATGTGTGATTTTACG 889
DB 878 ---CTAAGGTCTGTACAGCAAAAGGATGTTGATCCCACTGGCCACGGAGATGTTTTCC 934
QY 890 CCAGTTTCTACAACCTCTCGATTGCTTGATACCTTTATAGGAGAAAGCAAGAGTATATTT 949
DB 935 CCGCTCTTTTGAACAGTGGGAAGCTTGATGAATCTCTCACAGGTAAGGAGTAGCTGT 994
QY 950 TTGTGCTTAAACATAGATTAATCTGGGTGCCACAGTGGATCTGTATATTCTTAAATCATCTAA 1009
DB 995 TCATAGCTAACTCAGACAACTTTGGGAGCAATAGTTGATCTCAAAATTTTGAATCATCTGG 1054
QY 1010 TGAACCCCACTCAATGGAACCGCTGTGAATTTGTCTGAAAGCTCACAATAAAACACGCTG 1069
DB 1055 TCAA-----AAATAAATGAATCTGCTGAGGTCACTCCCAAAAGCTTTG 1102
QY 1070 CAGATGTAAGGGGGGACACTCACTCAATATGAAGCAAACTGAGACTGGTGGAAATTTG 1129
DB 1103 CAGATGTAAGAGTGGTACACTTATCTCTTATGAAGGCAGAGTTTCTAGAGATTG 1162
QY 1130 CTCAAGTGCCTCAAAAGCACATGTAGACGAGTTCAGTCTGTATCAAAAGTTCAAAATATTTA 1189
DB 1163 CACAGTTTCCAGAGAGAGCATGTTGGCGAATTCAGGCGATTTGAAAGTTTCAAAATCTTCA 1222
QY 1190 ATACAAACAACTATGGAATTTCTCTCTGAGCAGATTTAAAGACTGCAGAGCAAAATGCCA 1249
DB 1223 ATACCAATATTTATGGTGAACCTTGAAGGCGATTTAAGAGGCTTGTGAGGCTGATGCTC 1282
QY 1250 TTGACATGGAATCATTTGTAATGCAAGACTTTTGGATGGAGGCTGTAATGTCATTCAT 1309
DB 1283 TTAATAATGGAGATTAATCTCTAATCAAGGAAGTGGATGGG---TAAAGGTGCTTCAGC 1339
QY 1310 TAGAAATGCAGTAGGGGCTGCCATCAAAAGTTTGAAGATTTCTCTAGTATTATATGTC 1369
DB 1340 TTGAAACGGCTGCAGGTGCTGTATAAGGTTTGTGATCGTCAATTTGGTGTCAATGTTTC 1399
QY 1370 CAAGGAGCGCTTTCTGCTGTCAAAACCAACATCAGATCTCTTGTGCTGATGTCAAACC 1429
DB 1400 CAAGATCAGCTTTCTCCAGTGAAGGCACTTCAGATTTACTACTTGTGCGATCGGATC 1459
QY 1430 TCTATAGTCTTAATGACGAGTCTCTGACAATGAGTGAAGGCGGAATTTCTTACAGTGC 1489
DB 1460 TTTACACTGTGGAGGAAGGTTTGTGATCCGAAATCCCTGCTAGACTCAACCCCTACAAATC 1519
QY 1490 CCGTGGTTAAATTAGGCAGTCTTTTACGAAGTTCAAGATTTATCTAAGAAGATTTGAA 1549
DB 1520 CCACCATTTGAGTGGGCTCTGAAATTCAGAGAGTTTGGCAACTTTCTAAGCGTTTCAAGT 1579
QY 1550 GTATACCAAGATATGCTTCAATTTGGATCACCTCAGAGTTTTCAGGAGATGTGACATTTGGAA 1609
DB 1580 CCATACCTAGCATCATGATCTTGATAGTCTTAAGGTGTGAGTGTGTTGGTTGGCA 1639
QY 1610 AAAATGTTTCAAAAGGGAACGGTTTATCATCTTTCAGAAATCATGGTGAAGAATTTGATA 1669
DB 1640 GTGGAGTCACTCTGGAAGGGAAGGTTAATTAATTGAAGCAAGAGGCGGCTACACTTGAGA 1699
QY 1670 TCCCACTGGAGCAGTATTAGAGAACAAAGATTGTG 1704
DB 1700 TACCTGATGAATCTGTGATAGAAAACAAAGTAGTG 1734

Db 1568 CATTACTCTCAGGGAAGTGGACCATTCGTGCAAAACCTGGGGTTAAGTCGGAATTC 1627

QY 1674 ACCTGGAGCAGTATTAGAACAAAG 1698

Db 1628 GGATGGGTGGTCTGCGAAACAAAG 1652

RESULT 13

AAF94852/c

ID AAF94852 standard; cdNA; 396 BP.

XX AAF94852;

XX

XX 23-MAY-2001 (first entry)

XX Human ovarian cancer associated coding sequence SEQ ID NO: 43.

DE

XX Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.

XX

XX Homo sapiens.

XX

XX WO200118046-A2.

XX

XX 15-MAR-2001.

XX

XX 08-SEP-2000; 2000WO-US24827.

XX

XX 10-SEP-1999; 99US-0394374.

XX

XX 01-MAY-2000; 2000US-0561778.

XX

XX 15-AUG-2000; 2000US-0640173.

XX

XX 07-SEP-2000; 2000US-0656668.

XX

XX (CORI-) CORIXA CORP.

XX

XX Xu J, Stolk JA;

XX

XX WPI; 2001-211395/21.

XX

XX Isolated polypeptides associated with ovarian carcinomas, and the

XX

XX nucleic acids that encode them, useful for the prevention diagnosis and

XX

XX treatment of ovarian cancers -

XX

XX Claim 5; Page 129-130; 189pp; English.

XX

XX The present invention provides a number of coding sequences and proteins,

XX

XX the over-expression of which is associated with ovarian carcinoma/cancer.

XX

XX These can be used in the diagnosis, treatment and prevention of ovarian

XX

XX cancer, optionally by gene therapy or in the form of a vaccine. The

XX

XX present sequence is an example of one of these sequences.

XX

XX

SQ Sequence 396 BP; 125 A; 73 C; 51 G; 141 T; 6 other;

Query Match 17.5%; Score 375.2; DB 22; Length 396;

Best Local Similarity 97.7%; Pred. No. 6.2e-92;

Matches 377; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1684 GTATTAGAGAACAGATTGTGCTGGAACCTTCGCATCTTCGACACCTGAAATGAAAAA 1743

Db 396 GTATTAGAGAACAGATTGTGCTGGAACCTTCGCATCTTCGACACCTGAAATGAAAAA 337

QY 1744 TACTGTGGACACTTAAATGAGGCTAGTTCTTACAATGAAATGTTCTCTAGGATCTA 1803

Db 336 TACTGTGGACACTTAAATGAGGCTAGTTCTTACAATGAAATGTTCTCTAGGATCTA 277

QY 1804 AAATAGGAGGACTTTTACTATGTTACTGACCTGAGTGTGATTTTAAATAGACT 1863

Db 276 AAATAGGAGGACTTTTACTATGTTACTGACCTGAGTGTGATTTTAAATAGACT 217

QY 1864 TTTCTCAGTATGCTTTTACTCTAAGAAAGACAGATGGAGCAATACCTTCTCTTTG 1923

Db 216 TTTNCGATGATGCTTTTACTGTTNAGAAAGACAGATGGAGCAATACCTTCTCTTTG 157

QY 1924 AAGAGAAATCCAAAAGTTAGTTTCATCTTAAAGTGCATATTTGTTTAACTTAAACTGGG 1983

Db 156 AAGAGAAATCCAAAAGTTAGTTTCATCTTAAAGTGCATATTTTAAATTTAAACTGGG 97

QY 1984 CAACCTTTGGAAGAACTTTTAAACAGAGCCTCAATGATGATCACTTTTGAATTTGTTGAT 2043

Db 96 CAACCTTTGGAAGAACTTTTAAACAGAGCCTCAATGATGATCACTTTTGAATTTGTTGAT 37

QY 2044 TTCAAAAATAAAGCAGTGAAGCAATA 2069

Db 36 TTCAAAAATAAAGCAGTGAAGCAATA 11

RESULT 14

ABT03119/c

ID ABT03119 standard; cdNA; 396 BP.

XX

XX ABT03119;

XX

XX 05-SEP-2002 (first entry)

XX

XX Human ovarian carcinoma associated coding sequence SEQ ID NO: 43.

XX

XX Human; ovarian cancer; ovarian carcinoma; gene therapy; immunotherapy;

XX

XX cytostatic; gene; ss.

XX

XX Homo sapiens.

XX

XX WO200239885-A2.

XX

XX 23-MAY-2002.

XX

XX 13-NOV-2001; 2001WO-US45395.

XX

XX 14-NOV-2000; 2000US-0713550.

XX

XX 03-APR-2001; 2001US-0825294.

XX

XX 02-OCT-2001; 2001US-0970966.

XX

XX (CORI-) CORIXA CORP.

XX

XX Xu J, Stolk JA, Algate PA, Fling SP, Molesh DA;

XX

XX WPI; 2002-500186/53.

XX

XX Novel ovarian cancer polypeptide and polynucleotide, useful for

XX

XX detecting the presence of ovarian cancer in a patient, and in

XX

XX pharmaceutical compositions, e.g. vaccines, for treating ovarian cancer

XX

XX

XX Example 1; Page 127; 197pp; English.

XX

XX The present invention provides human ovarian cancer associated proteins

XX

XX and coding sequences. The sequences can be used in the diagnosis and

XX

XX treatment of ovarian cancers. The present sequence is a coding sequence

XX

XX of the invention.

XX

SQ Sequence 396 BP; 125 A; 73 C; 51 G; 141 T; 6 other;

Query Match 17.5%; Score 375.2; DB 24; Length 396;

Best Local Similarity 97.7%; Pred. No. 6.2e-92;

Matches 377; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1684 GTATTAGAGAACAGATTGTGCTGGAACCTTCGCATCTTCGACACCTGAAATGAAAAA 1743

Db 396 GTATTAGAGAACAGATTGTGCTGGAACCTTCGCATCTTCGACACCTGAAATGAAAAA 337

QY 1744 TACTGTGGACACTTAAATGAGGCTAGTTCTTACAATGAAATGTTCTCTAGGATCTA 1803

Db 336 TACTGTGGACACTTAAATGAGGCTAGTTCTTACAATGAAATGTTCTCTAGGATCTA 277

QY 1804 AAATAGGAGGACTTTTACTATGTTACTGACCTGAGTGTGATTTTAAATAGACT 1863

Db 276 AAATAGGAGGACTTTTACTATGTTACTGACCTGAGTGTGATTTTAAATAGACT 217

QY 1864 TTTCTGCAATGCTTTTAGTCTTAAGAAAGCAGATGGAGCAATCTTTCTTCTTTG 1923
Db |||||
QY 1924 AAGAGAAATCCCAAAAGTTAGTTTCACTTTAAAGTGCATATTTTAAATCTTAAACTGGG 1983
Db |||||
QY 1984 CAACCTTTGGAAGAACTTTTAAAGAGCCTCAATGATGATCACTTTGAAATGCTTTGAT 2043
Db |||||
QY 2044 TTCAAAAATAAAGCAGTGAAGCAATA 2069
Db |||||
36 TTCAAAAATAAAGCAGTGAAGCAATA 11

RESULT 15
ABL48802/c
ID ABL48802 standard; cDNA; 396 BP.
XX
AC ABL48802;
XX
DT 18-JUN-2002 (first entry)
XX
DE Ovarian carcinoma sequence isolate 24357.
XX
KW Ovarian cancer; cancer therapy; vaccine; gene therapy; tumour; cancer;
KW ss.
OS Homo sapiens.
XX
XX US2002004491-A1.
PN
PD 10-JAN-2002.
XX
XX 03-APR-2001; 2001US-0825294.
XX
PR 10-SEP-1999; 99US-0394374.
PR 01-MAY-2000; 2000US-0561778.
PR 15-AUG-2000; 2000US-0640173.
PR 07-SEP-2000; 2000US-0656668.
PR 14-NOV-2000; 2000US-0713550.
XX

(XUJ/) XU J.
(STOL/) STOLK J A.
(ALGA/) ALGATE P A.
(FLIN/) FLING S P.
PI Xu J, Stolk JA, Algate PA, Fling SP;
XX WPI; 2002-171027/22.
XX
PT Ovarian tumour polypeptide and polynucleotide useful in diagnosis,
PT prevention and/or treatment of cancer, especially ovarian cancer .
XX
XX Claim 1a; Page 53; 131pp; English.
XX

XX The invention relates to ovarian tumour polynucleotides and polypeptides
XX that may be utilised in cancer therapy, for example in a vaccine or
XX gene therapy. Polypeptides and polynucleotides of the invention are
XX useful for detecting a cancer in a patient, for stimulating and/or
XX expanding T-cells specific for a tumour protein, and for inhibiting the
XX development of a cancer in a patient. They are also useful for
XX stimulating an immune response in a patient, and for treating a cancer in
XX a patient and for determining the presence of a cancer in a patient.
XX The isolated polynucleotides of the invention are useful for their
XX ability to selectively form duplex molecules with complementary stretches
XX of the entire desired gene or gene fragments, and for designing and
XX preparing ribozyme molecules for inhibiting expression of tumour
XX polypeptides in tumour cells. Polypeptides and polynucleotides of the
XX invention are also useful in recombinant DNA molecules to direct
XX expression of a polypeptide in appropriate host cells. The sequences

CC Given in records ABL48760-ABL48956 represent polynucleotides encoding
CC ovarian carcinoma proteins.
XX
SQ Sequence 396 BP; 125 A; 73 C; 51 G; 141 T; 6 other;
Query Match 17.5%; Score 375.2; DB 24; Length 396;
Best Local Similarity 97.7%; Pred. No. 6.2e-92;
Matches 377; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1684 GTATTAGAGAACAGATTGTCTGGAAACCTTCGCATCTTGGACCACTGAAATGAAAA 1743
Db |||||
QY 1744 TACTGTGCACACTTAAATAATGGCTAGTTCTTACAATGAAATCTTCTAGGATTCTA 1803
Db |||||
QY 1804 AAATAGGCAGGTACTTTTACTATGTGTACCTGCAGTGTGATTTTTTAAATAGAGT 1863
Db |||||
QY 1864 TTTCTGCAATGCTTTTAGTCTTAAGAAAGCAGATGGAGCAATCTTTCTTCTTTG 1923
Db |||||
QY 1924 AAGAGAAATCCCAAAAGTTAGTTTCACTTTAAAGTGCATATTTTAAATCTTAAACTGGG 1983
Db |||||
QY 1984 CAACCTTTGGAAGAACTTTTAAAGAGCCTCAATGATGATCACTTTGAAATGCTTTGAT 2043
Db |||||
QY 2044 TTCAAAAATAAAGCAGTGAAGCAATA 2069
Db |||||
36 TTCAAAAATAAAGCAGTGAAGCAATA 11

Search completed: June 28, 2003, 02:58:54
Job time : 314.377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:39:45 ; Search time 71.5567 Seconds
(without alignments)
9201.581 Million cell updates/sec

Title: US-09-981-353-104

Perfect score: 2147

Sequence: 1 gtgtttaggagaagtaggg.....gttactatattagttggaag 2147

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	324.2	15.1	634	US-09-385-982-511	Sequence 511, App
C 2	70.2	3.3	3640	US-08-627-873-6	Sequence 6, Appl
C 3	57	2.7	7218	US-08-232-463-14	Sequence 14, Appl
C 4	43.2	2.0	304	US-08-627-873-4	Sequence 4, Appl
C 5	39.8	1.9	6265	US-09-129-112-3	Sequence 3, Appl
C 6	38.4	1.8	2205	US-08-840-767-53	Sequence 53, Appl
C 7	38.4	1.8	2449	US-08-840-767-7	Sequence 7, Appl
C 8	37.6	1.8	843	US-09-118-408-23	Sequence 23, Appl
C 9	37.6	1.8	843	US-09-506-855-23	Sequence 23, Appl
C 10	37.4	1.7	648	US-08-943-731-126	Sequence 126, App
C 11	37.4	1.7	24183	US-08-943-731-3	Sequence 3, Appl
C 12	37.2	1.7	12666	US-08-961-527-137	Sequence 137, App
C 13	37.2	1.7	53526	US-08-658-136-2	Sequence 2, Appl
C 14	37.2	1.7	53577	US-08-658-136-1	Sequence 1, Appl
C 15	36.4	1.7	1114	US-09-378-088A-109	Sequence 109, App
C 16	36.4	1.7	1152	US-08-844-188-37	Sequence 37, Appl
C 17	36.4	1.7	1152	US-09-378-088A-37	Sequence 37, Appl
C 18	36.4	1.7	1498	US-07-965-668A-1	Sequence 1, Appl
C 19	36.4	1.7	1498	US-08-950-433-1	Sequence 1, Appl
C 20	36.4	1.7	1498	US-09-186-287-1	Sequence 1, Appl
C 21	36.4	1.7	2230	US-08-844-188-34	Sequence 34, Appl
C 22	36.4	1.7	2230	US-09-378-088A-34	Sequence 34, Appl
C 23	35.8	1.7	2700	US-08-484-105-5	Sequence 5, Appl
C 24	35.8	1.7	2700	US-08-484-105-5	Sequence 5, Appl
C 25	35.8	1.7	4327	US-08-961-527-117	Sequence 117, App
C 26	35.8	1.7	4500	US-08-743-637B-35	Sequence 35, Appl
C 27	35.8	1.7	4500	US-08-526-840B-35	Sequence 35, Appl

28 35.8 1.7 43676 3 US-09-356-952-12 Sequence 12, Appl
C 29 35.6 1.7 2356 1 US-08-105-483-222 Sequence 222, App
C 30 35.6 1.7 2356 1 US-08-220-151-75 Sequence 75, Appl
C 31 35.6 1.7 2356 1 US-08-413-118-75 Sequence 75, Appl
C 32 35.6 1.7 2356 1 US-08-224-657-51 Sequence 51, Appl
C 33 35.6 1.7 2356 1 US-08-709-209-222 Sequence 222, App
C 34 35.6 1.7 2356 1 US-08-458-101-222 Sequence 222, App
C 35 35.6 1.7 2356 2 US-08-184-009-78 Sequence 78, Appl
C 36 35.6 1.7 2356 2 US-08-417-210A-68 Sequence 68, Appl
C 37 35.6 1.7 2356 2 US-08-458-356-78 Sequence 78, Appl
C 38 35.6 1.7 2356 3 US-08-473-446-75 Sequence 75, Appl
C 39 35.6 1.7 2356 4 US-08-460-736-78 Sequence 78, Appl
C 40 35.6 1.7 2356 4 US-09-354-138-51 Sequence 51, Appl
C 41 35.2 1.6 694 1 US-08-250-314-1 Sequence 1, Appl
C 42 35.2 1.6 694 1 US-08-708-107-1 Sequence 1, Appl
C 43 35 1.6 9510 4 US-09-453-702B-256 Sequence 256, App
C 44 35 1.6 98844 4 US-09-791-211-10 Sequence 10, Appl
C 45 34.4 1.6 36159 4 US-09-749-588-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-385-982-511/c
; Sequence 511, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDERGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCNDA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 511
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(634)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-511

Query Match 15.1%; Score 324.2; DB 4; Length 634;
Best Local Similarity 87.5%; Pred. No. 1.7e-81;
Matches 378; Conservative 0; Mismatches 50; Indels 4; Gaps 3;
QY 1398 CACATCAGATCTCTTGGTGGTATGCTCAACCTCTATAGTCTTAATGCAGGATCTCTGAC 1457
605 CAACCCACACAGTTTNTGNGNNAANNCACCCCTTAANNTTAAATGAGGNTCTGCC 546
1458 AATGAGTGAAGCGGGAAATTCCTACAGTGCCTTGGTTAAATAGGAGTCTTTTAC 1517
545 AATGA--GGAAAAGGGAAATTTCTTCNCAG--GGCCTGNTGNTTAATTAGCCAGTCTTTTAC 489
1518 GAAGTTTCAAGTATCTTAAGANGATTGGAAAGTANCCAGATATGCTTTGAATTTGGATC 1576
488 GAAGTTTCAAGTATCTTAAGANGATTGGAAAGTANCCAGATATGCTTTGAATTTGGATC 429
1577 ACCTCACAGTTTCAGGAGATGTGACATTTGGAAAAAATGTTTCATTAAAGGGAACGGTTA 1636
428 NCCTCNCAGTTTCAGGAGATGTGACATTTGNAAAAAATGTTTCATTAAAGGGAACGGTTA 369
1637 TCATCATTTGCAATCATGTCAGACAGATTGATATCCACCTGGAGCAGTATTAGAGAAC 1696

Db 368 TCATCATTCGCAATCATGGGACAGAAATTGATATCCACCTCGAGCAGTATTAGAGAACA 309
Qy 1697 AGATTGTGTCTGGAACCTTCGACCTCTGGACCACTGAAATGAAAAATCTGTGGACACT 1756
Db 308 AGATTGTGTCTGGAACCTTCGACCTCTGGACCACTGAAATGAAAAATCTGTGGACACT 249
Qy 1757 TAAATAATGGGCTAGTTCTTACAATGAAATGTTCTTAGGATTCTTAAATAGGCAGGTA 1816
Db 248 TAAATAATGGGCTAGTTCTTACAATGAAATGTTCTTAGGATTCTTAAATAGGCAGGTA 189
Qy 1817 CTTTACTATGTT 1828
Db 198 CTTTCTACTTT 177

RESULT 2
US-08-627-873-6
; Sequence 6, Application US/08627873
; Patent No. 5928932
; GENERAL INFORMATION:
; APPLICANT: Jarvis, Eric E.
; TITLE OF INVENTION: ISOLATED GENE ENCODING AN ENZYME WITH
; TITLE OF INVENTION: UDP-GLUCOSE PYROPHOSPHORYLASE AND
; TITLE OF INVENTION: PHOSPHOGLUCOMUTASE ACTIVITIES FROM
; TITLE OF INVENTION: CYCLOTELLA CRYPTICA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: National Renewable Energy Laboratory
; STREET: 1617 Cole Boulevard
; CITY: Golden
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80401-3393
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(DOS)text (*.*)
; CURRENT APPLICATION NUMBER: US/08/627,873
; FILING DATE: April 3, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Edna M. O'Connor
; REGISTRATION NUMBER: 29252
; REFERENCE/DOCKET NUMBER: 95-57
; TELEPHONE: 303/384-7573
; TELEFAX: 303/384-7499
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3640 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; FEATURE: 1.
; NAME/KEY: CDS (exon)
; LOCATION: 1..24
; NAME/KEY: Intron
; LOCATION: 25..314
; NAME/KEY: CDS (exon)
; LOCATION: 315..782
; NAME/KEY: Intron
; LOCATION: 783..885
; NAME/KEY: CDS (exon)
; LOCATION: 886..1402
; NAME/KEY: Intron
; LOCATION: 1403..1478
; NAME/KEY: CDS (exon)

; LOCATION: 1479..3637
US-08-627-873-6
Query Match 3.3%; Score 70.2; DB 2; Length 3640;
Best Local Similarity 54.4%; Pred. No. 8.7e-10;
Matches 141; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
Qy 498 GCCTGATAATATATCTTCCTCGGTGTTGAACAAACTAGTGTGTGTAACCTCAATGTGGTGT 557
Db 461 GGCTAGAGCCCATCTCTCCCTGCTCGCTTCGACGGTGTGTGAAGCTCAACGGTGGCCT 520
Qy 558 GGGAAACACGACATGGGCTGCAAAAGGCCCTTAAAGTCTGATTGGTGTGAGGAATGAGAATAC 617
Db 521 GGGTACGGGCATGGGTTGGATAAGGCTAAGAGTCTTTTGGAGGTGAAGAATGGAGATAC 580
Qy 618 CTTTCTGATCTGACTGTTTCAACAAATTTGAACATTTTGAATAAAACCTACAATACAGATGT 677
Db 581 CTTTGTGATTTGACGGCTAAGCAAGTCATGTGATGAGGGAGGAGTTTGGACAGAAAGT 640
Qy 678 TCCCTCTGTTTAAATGAATCTTTTAAACACGGATGAAGATACCAAAAAAATACTACAGAA 737
Db 641 CAAGTTTATGTGATGAACAGTTTTTCGACCTCGGATGATACTTTAGAGTTTTTTTAGGAC 700
Qy 738 GTACAATCATTTGCTGTGTG 756
Db 701 CAAGTATCTACTCTTTGG 719
RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOMLPX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Hardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-Fls
; US-08-232-463-14

Query Match
Best Local Similarity 2.7%; Score 57; DB 1; Length 7218;
Matches 15; Conservative 223; Mismatches 153; Indels 0; Gaps 0;

Qy 4 TTTAGGAGAAAGTAGGGCTGCTGGGTGCGGAGCGGCTGACGGGTGGCAAGGGGGG 63
Db 1442 TTTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1393

Qy 64 TTACAGCTGGGCTGCGACCCCTTAGGAGGGGCTCAAGGTGCGATGTCGAGGAAG 123
Db 1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323

Qy 124 AGAGAGAGAGGGCGGCTCAGAGTGACATTTTACGCTGCGAGCTTCTCCCGGGCG 183
Db 1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263

Qy 184 CCATAAGCGCCCAATTTCCAGCTGCTAAAGAGAGAGGAGATCTTAGCAAGCAATG 243
Db 1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203

Qy 244 TCTCAAGATGCTGCTTCTCAGTTCAGAGAGTCAATCGGCAAGAGCTAGATTCGTG 303
Db 1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143

Qy 304 AAGAAGGAACTAGAAAAAATACTACACAGCATCATCATGATTTGACGACACCAA 363
Db 1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1083

Qy 364 AAGACCTGGATGATTTTCGGAAGCTATTC 394
Db 1082 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1052
```

```
RESULT 4
US-08-627-873-4
; Sequence 4, Application US/08627873
; Patent No. 528932
; GENERAL INFORMATION:
; APPLICANT: Jarvis, Eric E.
; APPLICANT: Roessler, Paul G.
; TITLE OF INVENTION: ISOLATED GENE ENCODING AN ENZYME WITH
; TITLE OF INVENTION: UDP-GLUCOSE PYROPHOSPHORYLASE AND
; TITLE OF INVENTION: PHOSPHOGLUCOMUTASE ACTIVITIES FROM
; TITLE OF INVENTION: CYCLOTELLA CRYPTICA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: National Renewable Energy Laboratory
; STREET: 1617 Cole Boulevard
; CITY: Golden
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80401-3393
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(DOS)text (*.*)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,873
; FILING DATE: April 3, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Edna M. O'Connor
; REGISTRATION NUMBER: 29252
; REFERENCE/DOCKET NUMBER: 95-57
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/384-7573
; TELEFAX: 303/384-7499
; INFORMATION FOR SEQ ID NO: 4:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 304 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-627-873-4

Query Match
Best Local Similarity 2.0%; Score 43.2; DB 2; Length 304;
Matches 75; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 877 GGTGATATTTACGCCATTTCTTACAACCTCGGATGCTGTGATCTTTATAGGAGGC 936
Db 2 GGAGACTGTGACGCTGCCCTCATCGGCTCAGTCTCTCTGCGCTCTCTGCGAGGAGGA 61

Qy 937 AAAGAGTATATTTTGTGTCCTAACATAGATAATCTGGGTGCCACAGTGGATCTGTATAT 996
Db 62 TACAAGTACATGTTCTGCTCTAACTCTGACACCTTGGTGCCACCTTGACCTGAAATC 121

Qy 997 CTTAATCA 1004
Db 122 CTCACCCA 129

RESULT 5
US-09-129-112-3
; Sequence 3, Application US/09129112
; Patent No. 6465716
; GENERAL INFORMATION:
; APPLICANT: Etzler, Marilyn E.
; APPLICANT: Murphy, Judith B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A No. 6465716 Factor Binding Protein From Legume Roots
; FILE REFERENCE: 023070-079810US
; CURRENT APPLICATION NUMBER: US/09/129,112
; CURRENT FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 08/907,226
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6265
; TYPE: DNA
; ORGANISM: Dolichos biflorus
; FEATURE:
; OTHER INFORMATION: genomic sequence of NBP46 (DB46)
; NAME/KEY: exon
; LOCATION: (633)..(944)
; NAME/KEY: intron
; LOCATION: (945)..(1022)
; NAME/KEY: exon
; LOCATION: (1023)..(1151)
; NAME/KEY: intron
; LOCATION: (1152)..(1559)
; NAME/KEY: exon
; LOCATION: (1560)..(1616)
; NAME/KEY: intron
; LOCATION: (1617)..(1697)
; NAME/KEY: exon
; LOCATION: (1698)..(1790)
; US-09-129-112-3

Query Match
Best Local Similarity 1.9%; Score 39.8; DB 4; Length 6265;
Matches 107; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 1760 ATAATGGGCTAGTTTCTTACAATGAATGTTCTCTAGGATTTCTAAATAGGAGGTACTT 1819
Db 1940 ATATTTGCGTATATTTGTGAAAAAGAAAGTAGATTTTTCATTATGTTAAAGTATTT 1999

Qy 1820 TACTATGTTACTGTACCTGCGAGTGTGATTTTAAATAGAGTCTTCTGCGAGTATGCTT 1879
```


QY 568 ATGGCTGCAAGGCCCTAAAGTCTGATGGTGAGGAATGAGAACTACCTTCTGGAT 627
Db 382 AARGNCARAAAGGNSNATGGGNGCNGGARMGNTGYAARWSNCAYTAYGNGCN 441
QY 628 CTGACTGTTGACAAATTCGAATTAACAACTCAATACAGATGTTCTCTGTT 687
Db 442 TTYWSNGTNGNGNARARCCNATGCAYSNAAYCATYATYCARACNGTNATHITY 501
QY 688 TTAATGAACCTCTTTTAAACAGGATGAAGATACCAAAAAAATACTACAGAAAGTACAATCAT 747
Db 502 GAYACNGARTTGTNAAYTNTAYGAYCAVTTTAAATGTTTACNGGNAARTTAYTGY 561
QY 748 TGTGCTGGAATCTACACTTTCATCAACAGAGTACCCGAGGATTAATAAGAACT 807
Db 562 TAYTNCNGGNTTAYTNTTYSNYTNAAYGTCAYACNTGGAAYCARAARGACN 621
QY 808 TTACTTCTGTAGCAAGGACG 829
Db 622 TAYTNCAYATHATGAARAAYG 643

RESULT 9
US-09-506-855-23
; Sequence 23, Application US/09506855
; Patent No. 6448221
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Lasser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND
; FILE REFERENCE: 99-12
; CURRENT APPLICATION NUMBER: US/09/506,855
; CURRENT FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate nucleotide sequence encoding zsig37
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(843)
; OTHER INFORMATION: Each N is independently any nucleotide.
US-09-506-855-23

Query Match 1.8%; Score 37.6; DB 4; Length 843;
Best Local Similarity 28.6%; Pred. No. 0.61;
Matches 75; Conservative 46; Mismatches 141; Indels 0; Gaps 0;
QY 568 ATGGCTGCAAGGCCCTAAAGTCTGATGGTGAGGAATGAGAACTACCTTCTGGAT 627
Db 382 AARGNCARAAAGGNSNATGGGNGCNGGARMGNTGYAARWSNCAYTAYGNGCN 441
QY 628 CTGACTGTTGACAAATTCGAATTAACAACTCAATACAGATGTTCTCTGTT 687
Db 442 TTYWSNGTNGNGNARARCCNATGCAYSNAAYCATYATYCARACNGTNATHITY 501
QY 688 TTAATGAACCTCTTTTAAACAGGATGAAGATACCAAAAAAATACTACAGAAAGTACAATCAT 747
Db 502 GAYACNGARTTGTNAAYTNTAYGAYCAVTTTAAATGTTTACNGGNAARTTAYTGY 561
QY 748 TGTGCTGGAATCTACACTTTCATCAACAGAGTACCCGAGGATTAATAAGAACT 807
Db 562 TAYTNCNGGNTTAYTNTTYSNYTNAAYGTCAYACNTGGAAYCARAARGACN 621
QY 808 TTACTTCTGTAGCAAGGACG 829
Db 622 TAYTNCAYATHATGAARAAYG 643

RESULT 10
US-08-943-731-126/c
; Sequence 126, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; STREET: FLR.
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,731
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628
; FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DOYLE LEARY Ph.D., KATHRYN
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 126:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 648 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-943-731-126
Query Match 1.7%; Score 37.4; DB 4; Length 648;
Best Local Similarity 53.8%; Pred. No. 0.61;
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 1871 AGTATGCTTTTAGTCTAAGAAAAGCAGATGAGCAATACCTTCTCTTTGAACAGAA 1930
Db 375 AATATGAATTAACCTTCCATTTAAATCATATTTCTTAATTTTCTTACTATGAAACCT 316
QY 1931 TCCCAAAAGTTAGTTCATCTTAAAGTGAATATTTGTTTAATCTTAAACTGGGCAACTTT 1990
Db 315 TTGTAAAACACTAGTTCAITTTGAAAATGCTTCAATTATATATCTAAGTTTGTGACATTGA 256
QY 1991 GGAAGAACTTTTAAACAGAGCCT 2013

Db 255 GACATCACAACTGTCAAAAGACT 233

RESULT 11

US-08-943-731-3/c
; Sequence 3, Application US/08943731
; Patent No. 6265157

GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.

APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.

APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.

APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN

APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO

APPLICANT: ALA-KOKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING

ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666

CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.

STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; CITY: PHILADELPHIA

STATE: PA
; COUNTRY: USA

ZIP: 19103-7086
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,731
; FILING DATE: 03-OCT-1997

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628

FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:

NAME: DOYLE LEARY Ph.D., KATHRYN
; REGISTRATION NUMBER: 36,317

REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991

TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
; LENGTH: 24183 base pairs

TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-943-731-3

Query Match 1.7%; Score 37.4; DB 4; Length 24183;

Best Local Similarity 53.8%; Pred. No. 3.6; Mismatches 0; Gaps 0;

Matches 77; Conservative 0; Indels 66; Length 24183;

QY 1871 AGTATGCTTTTGTCTAAGAAAGCAGATGGAGCAATCTTTCCTTTTGAAGAGAA 1930

Db 9435 AATATGAATACCTTCCATTTAACTCATATTTCTTTAAATTTTCTTACTAAGAAACCT 9376

QY 1931 TCCCAAAGTATGTCATCTTAAAGTGAATATGTTTAACTTAAACCTGGCAACTTT 1990

Db 9375 TTGTAAGAACTAGTTCATTTGAAATGCTTCATTATATATCTTAAGTTTGGAGCATTTGA 9316

QY 1991 GGAAGAACTTTTAAACAGAAGCCT 2013

Db 9315 GACATCACAACTGTCAAAAGACT 9293

RESULT 12

US-08-961-527-137

; Sequence 137, Application US/08961527

Patent No. 6420135
; GENERAL INFORMATION:

APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue

CITY: Rockville
; STATE: Maryland

COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527
; FILING DATE:

CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

APPLICATION NUMBER:
; FILING DATE:

ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 137:

SEQUENCE CHARACTERISTICS:
; LENGTH: 12666 base pairs

TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: linear
; US-08-961-527-137

Query Match 1.7%; Score 37.2; DB 4; Length 12666;

Best Local Similarity 52.6%; Pred. No. 3;

Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 929 GAGAAGGCAAGACTATATTTTGTCTTAACATAGATATCTGGTGCCACAGTGCATC 988

Db 6525 GTGAATACAGCTATTTATCTTAATAGTGAATGATGATATTTGTCTACAATGAATA 6584

QY 989 TGTATATCTTAATCATCTAATGAACCCACCAATGAAACGCTGTGAATTTGTCTATCG 1048

Db 6585 GTTTAATTTAATAATTACTAAGCTGGCAGAAGTTGAAAAACGCTGTCAGTTAATGGCTG 6644

QY 1049 AAGTCACAAATAAAACACGTGCAGATGTAAAGG 1082

Db 6645 ATGAAATAGAAAAACACGTCAGCTGTAATGG 6678

RESULT 13

US-08-658-136-2/c

; Sequence 2, Application US/08658136

Patent No. 6071717
; GENERAL INFORMATION:

APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M

APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D

US-09-378-088A-109

	Query Match	1.7%;	Score 36.4;	DB 4;	Length 1114;
	Best Local Similarity	51.2;	Pred. No. 1.5;		
	Matches 85;	Conservative 0;	Mismatches 81;	Indels 0;	Gaps 0;
Qy	1968.	TAATCTTTAAACATCGGCAACTTTTGGAGAACTTTTAAACAGAGCCTCAATGATGATGCACCT	2027		
Db	1046	TGATCTCGGTAAGAGGTCACATTTGAAGTATCATTTCTGAAGTTTGGATTTTCATTACA	987		
Qy	2028	TTGAATTTGCTTGTGATTTCAAAAATAAAGCAGTGGAAGCAATCTTGTGTACACTGGTACT	2087		
Db	986	CTAATTTCCGAACCATTTATTCGATATAATTTCAAAGAAGTAATAGTGGGAATCCTTATA	927		
Qy	2088	TTATAATGCTTAACTATATAAATCGGTTTATTGTTGTATAGACAGTTACT	2133		
Db	926	GAATTCATGATTCAGTCAGTTGGATTCATCTCTCATCTCTGATTTGTTCT	881		

Search completed: June 28, 2003, 08:19:55
Job time : 73.5567 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 06:24:51 ; Search time 202.744 Seconds
(without alignments)
15719.871 Million cell updates/sec

Title: US-09-981-353-104

Perfect score: 2147
Sequence: 1 ggttttaggagaaagtaggg.....gttactatattagttggaag 2147

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2147	100.0	2147	9	US-09-981-353-104
2	2147	100.0	2222	9	US-09-981-353-104
3	1961.2	91.3	2118	9	US-10-158-646-11
4	1924	89.6	2277	9	US-10-158-646-10
5	1676.6	78.1	1823	10	US-09-880-107-3338
6	558.8	26.0	566	9	US-10-102-524-1334
7	432.8	20.2	437	9	US-09-918-995-14100
8	429	20.0	430	9	US-09-918-995-3419
9	423.2	19.7	471	9	US-09-918-995-35346
10	422	19.7	422	9	US-09-918-995-17154
11	421.4	19.6	433	9	US-09-918-995-36518
12	419.8	19.6	671	10	US-09-917-800A-877
13	388	18.0	439	9	US-09-918-995-8128
14	386.2	18.0	428	10	US-09-960-352-6825
15	375.2	17.5	396	9	US-09-970-966-43
16	375.2	17.5	396	10	US-09-825-294-43
17	372.4	17.3	406	9	US-09-918-995-4877
18	338.4	15.8	388	10	US-09-960-352-3813
19	324.2	15.1	634	9	US-09-871-161-511

20	294	13.7	520	9	US-10-102-524-191	Sequence 191, App
21	294	13.7	520	9	US-10-102-524-509	Sequence 509, App
22	259.4	12.1	301	10	US-09-960-352-13801	Sequence 13801, A
23	112.4	5.2	520	9	US-09-770-961-40	Sequence 40, Appl
24	110.4	5.1	274	10	US-09-783-590-11554	Sequence 11554, A
25	110	5.1	464	10	US-09-864-761-543	Sequence 543, App
26	90	4.2	281	10	US-09-923-876-3461	Sequence 3461, App
27	87.8	4.1	117	10	US-09-960-352-2989	Sequence 2989, App
28	83	3.9	83	10	US-09-864-761-17338	Sequence 17338, A
29	82.4	3.8	630	10	US-09-734-569-61	Sequence 61, Appl
30	69.4	3.2	288	10	US-09-294-093B-3367	Sequence 3367, App
31	67.4	3.1	139	10	US-09-960-352-3760	Sequence 3760, App
32	45	2.1	335913	9	US-09-754-853A-2	Sequence 2, Appl
33	45	2.1	335913	9	US-09-754-853A-3	Sequence 3, Appl
34	42.2	2.0	289	10	US-09-294-093B-198	Sequence 198, App
35	40.4	1.9	2207	9	US-10-097-065-120	Sequence 120, App
36	40.4	1.9	2674	9	US-10-097-065-15	Sequence 15, Appl
37	39.8	1.9	2000	9	US-09-938-842A-2973	Sequence 2973, App
38	39.8	1.9	6265	10	US-09-129-112-3	Sequence 3, Appl
39	39.6	1.8	530	10	US-09-864-761-12634	Sequence 12634, A
40	38.6	1.8	1442	10	US-09-731-872-63	Sequence 63, Appl
41	38.6	1.8	1659	9	US-10-001-835-3	Sequence 3, Appl
42	38.4	1.8	1065	10	US-09-804-682-33	Sequence 33, Appl
43	38.2	1.8	594	9	US-10-123-185-10	Sequence 10, Appl
44	38.2	1.8	2212	9	US-10-174-590-325	Sequence 325, App
45	38.2	1.8	2212	9	US-10-176-758-325	Sequence 325, App

ALIGNMENTS

RESULT 1

US-09-981-353-104
; Sequence 104, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 104
; LENGTH: 2147
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 2555628CB1
US-09-981-353-104

Query Match	100.0%	Score	2147;	DB	9;	Length	2147;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	2147;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	GTGTTTAGGAGAAAGTAGGGCTGTGGTGTCTGGAGCGGCTGACGGTGTGACAAAGGGG	60				
Db	1	GTGTTTAGGAGAAAGTAGGGCTGTGGTGTCTGGAGCGGCTGACGGTGTGACAAAGGGG	60				
QY	61	GGGTTTAGCAGCTGGGGCTGCGACCGTTAGGAGGGGCTCAAGGTGTGCATGTGTGAGGGAA	120				
Db	61	GGGTTTAGCAGCTGGGGCTGCGACCGTTAGGAGGGGCTCAAGGTGTGCATGTGTGAGGGAA	120				
QY	121	GAGAGAGAGAGAAAGGGCGCTCAGAGTGTCTTTCAGCTTCGAGCGCTTCTTCCCGGG	180				
Db	121	GAGAGAGAGAGAAAGGGCGCTCAGAGTGTCTTTCAGCTTCGAGCGCTTCTTCCCGGG	180				
QY	181	GGCCCATAAACGCCCCCAATTTCCAGCTGTCTTAAAGGAGAGAGATCTTAGCAAGCA	240				
Db	181	GGCCCATAAACGCCCCCAATTTCCAGCTGTCTTAAAGGAGAGAGATCTTAGCAAGCA	240				

```
QY 241 ATGCTCTCAAGATGGTGTCTCTCAGTTCCAAAGATCATTCGGCAAGAGCTAGAAATATCT 300
DB |||||
QY 241 ATGCTCTCAAGATGGTGTCTCTCAGTTCCAAAGATCATTCGGCAAGAGCTAGAAATATCT 300
DB |||||
QY 301 GTGAAGAAGAACTAGAAAAAATCTCACACAGCATCATCATGAAATTCGAGCACACC 360
DB |||||
QY 301 GTGAAGAAGAACTAGAAAAAATCTCACACAGCATCATCATGAAATTCGAGCACACC 360
DB |||||
QY 361 AAAAAAGACCTGGATGGATTTCCGAAGCTATTTTCATAGATTTTTCGAAGAAAGGGCCT 420
DB |||||
QY 361 AAAAAAGACCTGGATGGATTTCCGAAGCTATTTTCATAGATTTTTCGAAGAAAGGGCCT 420
DB |||||
QY 421 TCTGTGGATTGGGAAAAAATCAGAGACCCCTGGAAGATTTCGATTCAACCTATGAAAG 480
DB |||||
QY 421 TCTGTGGATTGGGAAAAAATCAGAGACCCCTGGAAGATTTCGATTCAACCTATGAAAG 480
DB |||||
QY 481 ATAAAGGCCAGGGCTTGCCTGATATATATCTCCGTGTTGAAACAAACTAGTGGTGGTG 540
DB |||||
QY 481 ATAAAGGCCAGGGCTTGCCTGATATATATCTCCGTGTTGAAACAAACTAGTGGTGGTG 540
DB |||||
QY 541 AAATCAATGGTGGTTTGGGAACAGCATGGGCTGCAAGGCCCTTAAAGTCTGATTGGT 600
DB |||||
QY 541 AAATCAATGGTGGTTTGGGAACAGCATGGGCTGCAAGGCCCTTAAAGTCTGATTGGT 600
DB |||||
QY 601 GTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTTCAGCAAAATGAAATTTGAATAA 660
DB |||||
QY 601 GTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTTCAGCAAAATGAAATTTGAATAA 660
DB |||||
QY 661 ACCTCAATACAGATGTTCTCTGTTTAAATGAACTCTTTTAAACCGGATGAAGATACC 720
DB |||||
QY 661 ACCTCAATACAGATGTTCTCTGTTTAAATGAACTCTTTTAAACCGGATGAAGATACC 720
DB |||||
QY 721 AAAAAAATCTACAGAAATGACAAATCTGCTGTGGAATCTACATCTTCAATCAAGC 780
DB |||||
QY 721 AAAAAAATCTACAGAAATGACAAATCTGCTGTGGAATCTACATCTTCAATCAAGC 780
DB |||||
QY 781 AGGTACCCGAGGATTAATAAGAACTCTTTTACTTCTGTAGCAAGGACGTGCTTACTCA 840
DB |||||
QY 781 AGGTACCCGAGGATTAATAAGAACTCTTTTACTTCTGTAGCAAGGACGTGCTTACTCA 840
DB |||||
QY 841 GGGGAAAAATACAGAACTGGTACCCCTCCAGGTGATGATTTACGCCAGTTTCTAC 900
DB |||||
QY 841 GGGGAAAAATACAGAACTGGTGGTACCCCTCCAGGTGATGATTTACGCCAGTTTCTAC 900
DB |||||
QY 901 AACTCTGGATTGCTGATACCTTTATAGGAAGGCAAGAGATATATTTTGTGCTAAC 960
DB |||||
QY 901 AACTCTGGATTGCTGATACCTTTATAGGAAGGCAAGAGATATATTTTGTGCTAAC 960
DB |||||
QY 961 ATAGATAAATCTGGTGCCACAGTGGATCTGTATATTTCTTAAATCAATGAACCCACC 1020
DB |||||
QY 961 ATAGATAAATCTGGTGCCACAGTGGATCTGTATATTTCTTAAATCAATGAACCCACC 1020
DB |||||
QY 1021 AATGAAAAACGCTGTGAAATTTGTGATGGAAGTACAAATTAATAAACAACGTCAGATGAA 1080
DB |||||
QY 1021 AATGAAAAACGCTGTGAAATTTGTGATGGAAGTACAAATTAATAAACAACGTCAGATGAA 1080
DB |||||
QY 1081 GCGGGACACTCACTCAATATGAGCAAACTGAGACTGGTGGAAATTCCTCAAGTGCCA 1140
DB |||||
QY 1081 GCGGGACACTCACTCAATATGAGCAAACTGAGACTGGTGGAAATTCCTCAAGTGCCA 1140
DB |||||
QY 1141 AAGACACATGTAGACAGTTCAAGTCTGTATCAAGTTCAAAATATTTAATAAACAAC 1200
DB |||||
QY 1141 AAGACACATGTAGACAGTTCAAGTCTGTATCAAGTTCAAAATATTTAATAAACAAC 1200
DB |||||
QY 1201 CTATGGATTCTCTTGCAGAGTTAAAGACTGCAGGACAAATGCCATTGACATGGA 1260
DB |||||
QY 1201 CTATGGATTCTCTTGCAGAGTTAAAGACTGCAGGACAAATGCCATTGACATGGA 1260
DB |||||
QY 1261 ATCATTTGTAATGCAAGACTTTGGATGGAGCCCTGAATGTCATTAATAGAACTGCA 1320
DB |||||
QY 1261 ATCATTTGTAATGCAAGACTTTGGATGGAGCCCTGAATGTCATTAATAGAACTGCA 1320
DB |||||
QY 1321 GTAGGGCTGCCATCAAAAGTTTGAAGATTTCTTAGGTATTAATGTGCCAAGGCCGT 1380
DB |||||
```

```
DB 1321 GTAGGGCTGCCATCAAAAGTTTGAAGATTTCTCTAGGTATTAATGTGCCAAGGCCGT 1380
QY 1381 TTTCTGCCCTGTCAAAACACATCAGATCTCTTCTGGTGTAGTGTCAAACTCTATAGTCTT 1440
DB |||||
QY 1381 TTTCTGCCCTGTCAAAACACATCAGATCTCTTCTGGTGTAGTGTCAAACTCTATAGTCTT 1440
DB |||||
QY 1441 AATGACAGATCTCTGACAAATGAGTGAAGCGGAAATTTCTTACAGTGCCTTGGTTAAA 1500
DB |||||
QY 1441 AATGACAGATCTCTGACAAATGAGTGAAGCGGAAATTTCTTACAGTGCCTTGGTTAAA 1500
DB |||||
QY 1501 TTAGCAGTTCTTTTACGAAGTTTCAAGATTATCTAAGAAGATTTTGAAGTATACCAGAT 1560
DB |||||
QY 1501 TTAGCAGTTCTTTTACGAAGTTTCAAGATTATCTAAGAAGATTTTGAAGTATACCAGAT 1560
DB |||||
QY 1561 ATGCTTGAATTTGGATCACCTCACAGTTTCAGAGATGTGACATTTGGAAAAAATGTTTCA 1620
DB |||||
QY 1561 ATGCTTGAATTTGGATCACCTCACAGTTTCAGAGATGTGACATTTGGAAAAAATGTTTCA 1620
DB |||||
QY 1621 TTAAGGGAACGGTTTATCATTTGCAAAATCATGCTGACAGAAATGTATCCACCTGGA 1680
DB |||||
QY 1621 TTAAGGGAACGGTTTATCATTTGCAAAATCATGCTGACAGAAATGTATCCACCTGGA 1680
DB |||||
QY 1681 GCAGTATTAGAGAAACAAGATTGTGCTGGAACCTTTCGCATCTTGGACCCTGAAATGAA 1740
DB |||||
QY 1681 GCAGTATTAGAGAAACAAGATTGTGCTGGAACCTTTCGCATCTTGGACCCTGAAATGAA 1740
DB |||||
QY 1741 AAATACTGTGGACACTTAAATTAATGGGCTAGTTCCTTCAATGAAATGTTCTTAGGATT 1800
DB |||||
QY 1741 AAATACTGTGGACACTTAAATTAATGGGCTAGTTCCTTCAATGAAATGTTCTTAGGATT 1800
DB |||||
QY 1801 CTAAATAGGAGGACTTAAATTAATGGGCTAGTTCCTTCAATGAAATGTTCTTAGGATT 1860
DB |||||
QY 1801 CTAAATAGGAGGACTTAAATTAATGGGCTAGTTCCTTCAATGAAATGTTCTTAGGATT 1860
DB |||||
QY 1861 AGTTTCTGACAGTATGCTTTTAGTCTAAGAAAGCACAGATGGAGCAATCTTTCTTCT 1920
DB |||||
QY 1861 AGTTTCTGACAGTATGCTTTTAGTCTAAGAAAGCACAGATGGAGCAATCTTTCTTCT 1920
DB |||||
QY 1921 TTGAAGAAATCCCAAAAGTTAGTTTCATCTTAAAGTGAATATTTGTTAACTTTAAACT 1980
DB |||||
QY 1921 TTGAAGAAATCCCAAAAGTTAGTTTCATCTTAAAGTGAATATTTGTTAACTTTAAACT 1980
DB |||||
QY 1981 GGGCAACTTTTGAAGAACTTTTAAAGAGGCTCAATGATGATCACTTTGAATGCTTGT 2040
DB |||||
QY 1981 GGGCAACTTTTGAAGAACTTTTAAAGAGGCTCAATGATGATCACTTTGAATGCTTGT 2040
DB |||||
QY 2041 GATTTCAAAAAATAAAGCAGTGAAGCAATATCTTGTGTACACTGGTACTTTTAAATGCTAAC 2100
DB |||||
QY 2041 GATTTCAAAAAATAAAGCAGTGAAGCAATATCTTGTGTACACTGGTACTTTTAAATGCTAAC 2100
DB |||||
QY 2101 TATAAATGCTTTTATTTGTTGTAGACAGTACTATATTAGTTGGAAG 2147
DB |||||
QY 2101 TATAAATGCTTTTATTTGTTGTAGACAGTACTATATTAGTTGGAAG 2147
DB |||||
```

RESULT 2

```
US-09-981-353-76
; Sequence 76, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jorse, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 76
; LENGTH: 2222
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 237563.31
; NAME/KEY: unsure
; LOCATION: 2208
; OTHER INFORMATION: a, t, c, g, or other
; US-09-981-353-76

Query Match      100.0%; Score 2147; DB 9; Length 2222;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGTTTAGGAGAAAGTAGGGCTGTGGGTGTCGGAGCCGGCTGACGGGTGACAAAGGG 60
DB 46 GTGTTTAGGAGAAAGTAGGGCTGTGGGTGTCGGAGCCGGCTGACGGGTGACAAAGGG 105
QY 61 GGGTTAGCAGCTGGGCTGGACCGTTAGGAGGGGCTCAAGGTGTCATGTGTAGGGAA 120
DB 106 GGGTTAGCAGCTGGGCTGGACCGTTAGGAGGGGCTCAAGGTGTCATGTGTAGGGAA 165
QY 121 GAGAGAGAGAGAAAGGGCGCTCAGAGGTGACTTTTCAGCCTGCGAGCCCTTTCCCGGG 180
DB 166 GAGAGAGAGAGAAAGGGCGCTCAGAGGTGACTTTTCAGCCTGCGAGCCCTTTCCCGGG 225
QY 181 GCGCCATAAACGCCCCCAATTTCCAGCTGCTTAAGGAGAGAGAGATCTTAGCAAGCA 240
DB 226 GCGCCATAAACGCCCCCAATTTCCAGCTGCTTAAGGAGAGAGAGATCTTAGCAAGCA 285
QY 241 ATGTCTCAAGATGGTCTCTCAGTCCAGAGTCAATTCGGCAAGAGCTAGAAATATCT 300
DB 286 ATGTCTCAAGATGGTCTCTCAGTCCAGAGTCAATTCGGCAAGAGCTAGAAATATCT 345
QY 301 GTGAAGAGAGAACTAGAAAAATACCTCACACAGCATCATCATGAAATTCAGACACC 360
DB 346 GTGAAGAGAGAACTAGAAAAATACCTCACACAGCATCATCATGAAATTCAGACACC 405
QY 361 AAAAAAGACTGGATGGATTTGGAAGCTATTTATAGATATTTTGAAGAAAGGGGCT 420
DB 406 AAAAAAGACTGGATGGATTTGGAAGCTATTTATAGATATTTTGAAGAAAGGGGCT 465
QY 421 TCTGTGGATTTGGGAAAAATCAGAGACCCCTGGAAGATTCGATCAACCCCTATGAAG 480
DB 466 TCTGTGGATTTGGGAAAAATCAGAGACCCCTGGAAGATTCGATCAACCCCTATGAAG 525
QY 481 ATAAAGCCAGGGCTTGCTGATATATATCTTCGCTGTGAACAAACTAGTGTGGTG 540
DB 526 ATAAAGCCAGGGCTTGCTGATATATATCTTCGCTGTGAACAAACTAGTGTGGTG 585
QY 541 AAATCAATGGTGGTTGGAAACAGCATGGGCTGCAAGGCCCTAAAGTCTGATGGT 600
DB 586 AAATCAATGGTGGTTGGAAACAGCATGGGCTGCAAGGCCCTAAAGTCTGATGGT 645
QY 601 GTGAGGAATGAGNATACCTTTCTGGATCTGACTGTTCAGCAAAATGGAATTTGAATAA 660
DB 646 GTGAGGAATGAGNATACCTTTCTGGATCTGACTGTTCAGCAAAATGGAATTTGAATAA 705
QY 661 ACCTACAATACAGATGTTCTCTGTTTAAATGAATCTTTTAAACAGGATGAAGATCC 720
DB 706 ACCTACAATACAGATGTTCTCTGTTTAAATGAATCTTTTAAACAGGATGAAGATCC 765
QY 721 AAAAAAATACAGAGTACATCATTTGCTGTGAAATCTACATTTCAATCAAAAGC 780
DB 766 AAAAAAATACAGAGTACATCATTTGCTGTGAAATCTACATTTCAATCAAAAGC 825
QY 781 AGGTACCCGAGNATTAAGAAATCTTTACTTCTGTAGCAAGAGAGCTGTCTTACTCA 840
DB 826 AGGTACCCGAGNATTAAGAAATCTTTACTTCTGTAGCAAGAGAGCTGTCTTACTCA 885
QY 841 GGGGAAAAATACAGAGCTTTGGTACCTCCAGGTTCATGGTGATATTTTACGCCAGTTTCTAC 900
DB 886 GGGGAAAAATACAGAGCTTTGGTACCTCCAGGTTCATGGTGATATTTTACGCCAGTTTCTAC 945
QY 901 AACTCTGGATTCCTTGATACCTTTTATAGGAGAGGCAAGAGTATATTTTGTGTCTAAC 960
```

```

946 AACTCTGGATTCCTTGATACCTTTTATAGGAGAGGCAAGAGTATATTTTGTGTCTAAC 1005
QY 961 ATAGATAATCTGGGTGCCACAGTGATCTGTATATTTCTTAATCATCTAATGAACCCACC 1020
DB 1006 ATAGATAATCTGGGTGCCACAGTGATCTGTATATTTCTTAATCATCTAATGAACCCACC 1065
QY 1021 AATGAAACCGCTGTGAATTTTCTCATGGAAGTCACAAATATAAACACGTCAGATGTAAAG 1080
DB 1066 AATGAAACCGCTGTGAATTTTCTCATGGAAGTCACAAATATAAACACGTCAGATGTAAAG 1125
QY 1081 GCGGGACACCTCCTCAATATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCA 1140
DB 1126 GCGGGACACCTCCTCAATATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCA 1185
QY 1141 AAGGACACATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAATATTTAATCAACAAC 1200
DB 1186 AAGGACACATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAATATTTAATCAACAAC 1245
QY 1201 CTATGGATTTCTTTGCGAGCAGTTAAAAGACTGCGAGGAGCAAAATGCCAATTCACATGGAA 1260
DB 1246 CTATGGATTTCTTTGCGAGCAGTTAAAAGACTGCGAGGAGCAAAATGCCAATTCACATGGAA 1305
QY 1261 ATCATTTGTAATGCAAAAGCTTTTGGATGGAGGCTGAATGTCTCAATTTAGAAACTGCA 1320
DB 1306 ATCATTTGTAATGCAAAAGCTTTTGGATGGAGGCTGAATGTCTCAATTTAGAAACTGCA 1365
QY 1321 GTAGGGGCTGCCATCAAAAGTTTTCAGAAATTTCTCTAGGTATTAATGTGCGCAAGGAGCCGT 1380
DB 1366 GTAGGGGCTGCCATCAAAAGTTTTCAGAAATTTCTCTAGGTATTAATGTGCGCAAGGAGCCGT 1425
QY 1381 TTTCTGCTGTCAAAACCAACATCAGATCTCTTGTGTTGATGTCAAACTCTATAGTCTT 1440
DB 1426 TTTCTGCTGTCAAAACCAACATCAGATCTCTTGTGTTGATGTCAAACTCTATAGTCTT 1485
QY 1441 AATCAGGATCTCTGCAATGAGTGAAAGCGGGAATTTCTCTACAGTGCCCTTGGTTAAA 1500
DB 1486 AATCAGGATCTCTGCAATGAGTGAAAGCGGGAATTTCTCTACAGTGCCCTTGGTTAAA 1545
QY 1501 TTAGGACGTTCTTTTACGAAGTTTCAAGATTTATCTAAGAAATTTTGAAGTATACAGAT 1560
DB 1546 TTAGGACGTTCTTTTACGAAGTTTCAAGATTTATCTAAGAAATTTTGAAGTATACAGAT 1605
QY 1561 ATGCTTGAATTTGGATCACCTCACAGTTTTCAGGAGATGTGACATTTTGGAAAAATGTTTCA 1620
DB 1606 ATGCTTGAATTTGGATCACCTCACAGTTTTCAGGAGATGTGACATTTTGGAAAAATGTTTCA 1665
QY 1621 TTAAAGGAAACGGTTTATCATTTGCAAAATCATGTGTGACAGAAATTTGATATCCACCTGGA 1680
DB 1666 TTAAAGGAAACGGTTTATCATTTGCAAAATCATGTGTGACAGAAATTTGATATCCACCTGGA 1725
QY 1681 GCAGTATTTAGAGAACAGATTTGTGTCTGGAACCTTCCGATCTTGGACCACTGAAATGAA 1740
DB 1726 GCAGTATTTAGAGAACAGATTTGTGTCTGGAACCTTCCGATCTTGGACCACTGAAATGAA 1785
QY 1741 AAATACTGTGACACTTAAATAATGGGCTAGTTCTTCAAAATGAAATTTCTCTAGGATTT 1800
DB 1786 AAATACTGTGACACTTAAATAATGGGCTAGTTCTTCAAAATGAAATTTCTCTAGGATTT 1845
QY 1801 CTAAAAATAGGAGGTAATTTACTATGTTACTGTACCCCTGCAGTGTGATTTTAAAAATAG 1860
DB 1846 CTAAAAATAGGAGGTAATTTACTATGTTACTGTACCCCTGCAGTGTGATTTTAAAAATAG 1905
QY 1861 AGTTTCTGCGAGTATGCTTTTAGTCTAAGAAAGCAGATGGAGCAATACTTTCTCTCT 1920
DB 1906 AGTTTCTGCGAGTATGCTTTTAGTCTAAGAAAGCAGATGGAGCAATACTTTCTCTCT 1965
QY 1921 TTGAAGAGAAATCCCAAAAGTTAGTTTCATCTTAAAGTGAATATTTCTTAACTTAAACT 1980
DB 1966 TTGAAGAGAAATCCCAAAAGTTAGTTTCATCTTAAAGTGAATATTTCTTAACTTAAACT 2025
QY 1981 GGGCAACTTTTGGAGAACTTTTAAACAGAGGCTCAATGATGATCACTTTGAAATGCTTGT 2040
```

```
Db 2026 GGGCACTTTGGAAGACCTTTTAAACAGAGCCCTCAATGATGATCACTTTGAATTCCTTGT 2085
Qy 2041 GATTTCAAAATAAAGCAGTGAAGCAATACTTGTGTACACTGGTACTTTTATTAATGCTAAC 2100
Db 2086 GATTTCAAAATAAAGCAGTGAAGCAATACTTGTGTACACTGGTACTTTTATTAATGCTAAC 2145
Qy 2101 TATAAAGCTGTTTATTTGTTTGTAGACAGTACTATATTAATGTTGGAAG 2147
Db 2146 TATAAAGCTGTTTATTTGTTTGTAGACAGTACTATATTAATGTTGGAAG 2192

RESULT 3
US-10-158-646-11
; Sequence 11, Application US/10158646
; Publication No. US20030073105A1
; GENERAL INFORMATION:
; APPLICANT: Sornasse, Thierry
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0030-1 US
; CURRENT APPLICATION NUMBER: US/10/158,646
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/295,239
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 2118
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030073105A1 237563.10
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2110
; OTHER INFORMATION: a, t, c, g, or other
US-10-158-646-11

Query Match 91.3%; Score 1961.2; DB 9; Length 2118;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 2045; Conservative 0; Mismatches 4; Indels 62; Gaps 2;

Qy 4 TTTAGAGAAAGTAGGGGCTGTGGTGTGGGAGCGGCTGACGGGTGACAAAGGGGGG 63
Db 8 TTTAGAGAAAGTAGGGGCTGTGGTGTGGGAGCGGCTGACGGGTGACAAAGGGGGG 67
Qy 64 TTAGCAGCTGGGCTGCGACCGTTAGGGAGGGGCTCAAGGTGTGCATGTGTGAGGGGAAG 123
Db 68 TTAGCAGCTGGGCTGCGACCGTTAGGGAGGGGCTCAAGGTGTGCATGTGTGAGGGGAAG 127
Qy 124 AGAGAGAGAAAGGGCGCTCAGAGTGACTTTTCCAGCTGCGAGCCCTTCTTCCCGGGGG 183
Db 128 AGAGAGAGAAAGGGCGCTCAGAGTGACTTTTCCAGCTGCGAGCCCTTCTTCCCGGGGG 187
Qy 184 CCATTAACCCGCCAATTTCCAGCTGCTTAAGAGAGAGA----- 224
Db 188 CCATTAACCCGCCAATTTCCAGCTGCTTAAGAGAGAGAAGGTACCTGTGCGTGCACG 247
Qy 225 -----AGATCTTTAGCAAAAGCAAT 242
Db 248 CAGACGGGAAGGCTGGGNAAGCGGAGGACTGAGAAAGCCAGATCTTAGCAAGCAAT 307
Qy 243 GTCTCAAGATGGTGTCTCAGTTCGAAGAGTCAATTCGGCAAGAGCTAGATTAATCTGT 302
Db 308 GTCTCAAGATGGTGTCTCAGTTCGAAGAGTCAATTCGGCAAGAGCTAGATTAATCTGT 367
Qy 303 GAGAGAGAACTAGAAAATATCTCACCACAGCATCATCATGAATTTGAGCACACCAA 362
Db 368 GAAAGAGAACTAGAAAATATCTCACCACAGCATCATCATGAATTTGAGCACACCAA 427
Qy 363 AAAAGAGCTGATGAGTTTCGGAAGCTATTTTCATAGATTTTTCGAAAGAAAGGGGCTTC 422
Db
```

QY 1502 TAGGCAGTTCCTTTTACGAAGTTCAAGATTATCTAAGAAGATTGTGAAGTATACCAGATA 1561
DB 1568 TAGGCAGTTCCTTTTACGAAGTTCAAGATTATCTAAGAAGATTGTGAAGTATACCAGATA 1627
QY 1562 TGCCTGAATGGATCACCTCACAGTTTCAGGAGATGTGACATTTGGAAAAATGTTTCAT 1621
DB 1628 TGCCTGAATGGATCACCTCACAGTTTCAGGAGATGTGACATTTGGAAAAATGTTTCAT 1687
QY 1622 TAAAGGGAACGGTTATCATCTATTGCAATCATGGTGACAGAAATGATATCCCACTGGAG 1681
DB 1688 TAAAGGGAACGGTTATCATCTATTGCAATCATGGTGACAGAAATGATATCCCACTGGAG 1747
QY 1682 CAGTATTAGAGAACAAAGATTGTGTCTGGAAACCTTGGCAATCTTGGACCACTGAAATGAAA 1741
DB 1748 CAGTATTAGAGAACAAAGATTGTGTCTGGAAACCTTGGCAATCTTGGACCACTGAAATGAAA 1807
QY 1742 AATACCTGGACACACTTAATTAATGGCTAGTTTCTTACAAATGTTTCTTAGAATTC 1801
DB 1808 AATACCTGGACACACTTAATTAATGGCTAGTTTCTTACAAATGTTTCTTAGAATTC 1867
QY 1802 TAAATAGGCAGTACTTCTTACTATGTTACTGTACCTCGCAGTGTGATTTTAAAAATAGA 1861
DB 1868 TAAATAGGCAGTACTTCTTACTATGTTACTGTACCTCGCAGTGTGATTTTAAAAATAGA 1927
QY 1862 GTTTTCTGCAGTATGCTTTTACTCTAAGAAAAAGCACAGATGGAGCAATACTTTCTCTT 1921
DB 1928 GTTTTCTGCAGTATGCTTTTACTCTAAGAAAAAGCACAGATGGAGCAATACTTTCTCTT 1987
QY 1922 TGAAGAGATCCCAAGATTAGTTCATCTTAAAGTGCATATGTTTAAATCTTAAACTG 1981
DB 1988 TGAAGAGATCCCAAGATTAGTTCATCTTAAAGTGCATATGTTTAAATCTTAAACTG 2047
QY 1982 GCAACTTTGGAGAACTTTTAAAGAACCTTAAAGAACCTTAAAGAACCTTAAAGAACCTT 2041
DB 2048 GCAACTTTGGAGAACTTTTAAAGAACCTTAAAGAACCTTAAAGAACCTTAAAGAACCTT 2107
QY 2042 ATTTCAAAAAAT 2052
DB 2108 ATNTCAAAAAAT 2118

RESULT 4

US-10-158-646-10
; Sequence 10, Application US/10158646
; Publication No. US20030073105A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0030-1 US
; CURRENT APPLICATION NUMBER: US/10/158,646
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/295,239
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 2277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inbyte ID No. US20030073105A1 237563.4
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2263
; OTHER INFORMATION: a, t, c, g, or other

US-10-158-646-10

Query Match 89.6%; Score 1924; DB 9; Length 2277;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1924; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 AAGATCTTTAGCAAAAGCAATGTCTCAAGATGGTGTCTTCTCAGTTCCAAGAAGTCATTTCGCG 283
DB 324 AAGATCTTTAGCAAAAGCAATGTCTCAAGATGGTGTCTTCTCAGTTCCAAGAAGTCATTTCGCG 383
QY 284 AAGAGCTAGAAATTTATCTGTGAAGAAGAACTAGAAAAAATACTACCAAGATCATCATAC 343
DB 384 AAGAGCTAGAAATTTATCTGTGAAGAAGAACTAGAAAAAATACTACCAAGATCATCATAC 443
QY 344 ATGAATTTGAGCACACCAAAAGAACCTTGGATGGATTTTCGGAAGCTATTTTCATAGATTTT 403
DB 444 ATGAATTTGAGCACACCAAAAGAACCTTGGATGGATTTTCGGAAGCTATTTTCATAGATTTT 503
QY 404 TGCAGAAAAAGGGGCTTCTGTGGATTTGGGAAAAAATCCAGAGACCCCTCGAAGATTGCA 463
DB 504 TGCAGAAAAAGGGGCTTCTGTGGATTTGGGAAAAAATCCAGAGACCCCTCGAAGATTGCA 563
QY 464 TTCAACCTTATGAAGAAGATAAAGGCCAGGGGCTTGCCTGTATATATATCTTCGCTGTGCA 523
DB 564 TTCAACCTTATGAAGAAGATAAAGGCCAGGGGCTTGCCTGTATATATATCTTCGCTGTGCA 623
QY 524 ACAGAACTAGTGTGTGAACTCAATGTGTGTTTGGGAACCAAGCATGGGCTCAGAAAGGCC 583
DB 624 ACAGAACTAGTGTGTGAACTCAATGTGTGTTTGGGAACCAAGCATGGGCTCAGAAAGGCC 683
QY 584 CTAAAGTCTGATTTGGTGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTCAGCAAA 643
DB 684 CTAAAGTCTGATTTGGTGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTCAGCAAA 743
QY 644 TTGAACATTTGAATAAAAACTTCAATACAGATGTTCTCTGTTTAAATCTCTTTTAA 703
DB 744 TTGAACATTTGAATAAAAACTTCAATACAGATGTTCTCTGTTTAAATCTCTTTTAA 803
QY 704 ACACGGATGAGATACCAAAAAAATACTACAGAACTAGCAATCATTTCTGTGTGAAATCT 763
DB 804 ACACGGATGAGATACCAAAAAAATACTACAGAACTAGCAATCATTTCTGTGTGAAATCT 863
QY 764 ACACCTTTCAATCAAAAGCAGGTACCCGAGGATTAAATAAGAAATCTTTTACTTCTGTAGCAA 823
DB 864 ACACCTTTCAATCAAAAGCAGGTACCCGAGGATTAAATAAGAAATCTTTTACTTCTGTAGCAA 923
QY 824 AGGACGTGTCTTCTACAGGGGAAAAATACAGAACTGTGTACCCCTCAAGTCTATGTGTATA 883
DB 924 AGGACGTGTCTTCTACAGGGGAAAAATACAGAACTGTGTACCCCTCAAGTCTATGTGTATA 983
QY 884 TTTACGCCAGTCTTCTACAACTCTGATTTGATACCTTTTATAGGAGAACCAAGACT 943
DB 984 TTTACGCCAGTCTTCTACAACTCTGATTTGATACCTTTTATAGGAGAACCAAGACT 1043
QY 944 ATATTTTGTGTCTAAACATAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATC 1003
DB 1044 ATATTTTGTGTCTAAACATAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATC 1103
QY 1004 ATCTAATGAACCCACCCCAATGGAAAAACGCTGTGAATTTGTCTATGGAAGTCAAAATAAAA 1063
DB 1104 ATCTAATGAACCCACCCCAATGGAAAAACGCTGTGAATTTGTCTATGGAAGTCAAAATAAAA 1163
QY 1064 CACGTGCAGATGTAAAGGGGGGACACTCACTCAATATGAAGGCAAACTGACACTGGTGG 1123
DB 1164 CACGTGCAGATGTAAAGGGGGGACACTCACTCAATATGAAGGCAAACTGACACTGGTGG 1223
QY 1124 AAATTTGCTCAAGTGCACAAAGCACATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAA 1183
DB 1224 AAATTTGCTCAAGTGCACAAAGCACATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAAA 1283
QY 1184 TATTTAATACAAACCACTATGGATTTCTCTTTCAGAGAGTGTAAAAAGACTGCGAGGACAAA 1243
DB 1284 TATTTAATACAAACCACTATGGATTTCTCTTTCAGAGAGTGTAAAAAGACTGCGAGGACAAA 1343
QY 1244 ATGCCATTTGACATGGAATCATTTGTAATGCAAAAGACTTTTGGATGGAGCCCTGAATGTCA 1303
DB 1344 ATGCCATTTGACATGGAATCATTTGTAATGCAAAAGACTTTTGGATGGAGCCCTGAATGTCA 1403

1304 TTCAATTAGAACTGCAGTAGGGCTGCCATCAAAAGTTTTCAGAAATCTCTAGGTATTA 1363
1404 TTCAATTAGAACTGCAGTAGGGCTGCCATCAAAAGTTTTCAGAAATCTCTAGGTATTA 1463
1364 ATGTGCCAAGAGCGGTTTCTGCTGTGCTCAAAACACATCAGATCTCTTGTGGTGATG 1423
1464 ATGTGCCAAGAGCGGTTTCTGCTGTGCTCAAAACACATCAGATCTCTTGTGGTGATG 1523
1424 CAACCTCTATAGTCTTAATGAGGATCTCTGACATGAGTGAAGAGCGGATTTCTTA 1483
1524 CAACCTCTATAGTCTTAATGAGGATCTCTGACATGAGTGAAGAGCGGATTTCTTA 1583
1484 CAGTGCCTTGTGTTAAATAGGAGTCTTTTACGAAGGTTCAAGATTAATCTAAGAAGAT 1543
1584 CAGTGCCTTGTGTTAAATAGGAGTCTTTTACGAAGGTTCAAGATTAATCTAAGAAGAT 1643
1544 TTGAAGATATACAGATATGCTTGAATGAGTCACTCAGTTTCAGGAGATGTGACAT 1603
1644 TTGAAGATATACAGATATGCTTGAATGAGTCACTCAGTTTCAGGAGATGTGACAT 1703
1604 TTGGAAGAAATGTTTCATTTAAAGGNAACGTTTATCATCTGCAATCATGTGACAGAA 1663
1704 TTGGAAGAAATGTTTCATTTAAAGGNAACGTTTATCATCTGCAATCATGTGACAGAA 1763
1664 TTGATATCCACCTGGAGCAGTATTAGAGAAACAAGATTGCTGTGAAACCTTCGCATCT 1723
1764 TTGATATCCACCTGGAGCAGTATTAGAGAAACAAGATTGCTGTGAAACCTTCGCATCT 1823
1724 TGGACCACTGAAATGAAATATCTGTGGAACCTTAAATTAATGGGCTAGTTTCTTACAATG 1783
1824 TGGACCACTGAAATGAAATATCTGTGGAACCTTAAATTAATGGGCTAGTTTCTTACAATG 1883
1784 AAATGTTCTTAGGATTTTAAATAGGCAAGTACTTTACTATGTTTACTGTACCTGCAGT 1843
1884 AAATGTTCTTAGGATTTTAAATAGGCAAGTACTTTACTATGTTTACTGTACCTGCAGT 1943
1844 GTTGATTTTAAATAGAGTTTCTGCAGTATGCTTTTAACTTAAAGAAAGCAAGATGG 1903
1944 GTTGATTTTAAATAGAGTTTCTGCAGTATGCTTTTAACTTAAAGAAAGCAAGATGG 2003
1904 AGCAATACCTTCTTCTTTGAAGAGAAATCCAAAAGTTAGTTTCAATCTTAAAGTGCATAT 1963
2004 AGCAATACCTTCTTCTTTGAAGAGAAATCCAAAAGTTAGTTTCAATCTTAAAGTGCATAT 2063
1964 TGTGTTAATCTTAAATAGGCACTTTTGAAGAGAAATCCAAAAGTTAGTTTCAATCTTAAAGTGCATAT 2023
2064 TGTGTTAATCTTAAATAGGCACTTTTGAAGAGAAATCCAAAAGTTAGTTTCAATCTTAAAGTGCATAT 2123
2024 CACTTTGAATGCTTGTGATTTTCAAAATTAAGCAGTGAAGCAATCTTGTGTACACTGG 2083
2124 CACTTTGAATGCTTGTGATTTTCAAAATTAAGCAGTGAAGCAATCTTGTGTACACTGG 2183
2084 TACTTTAATGCTAATACTATAAAGTCTGTTTATTTGTTTGTAGACAGTTACTATATTAGTTG 2143
2184 TACTTTAATGCTAATACTATAAAGTCTGTTTATTTGTTTGTAGACAGTTACTATATTAGTTG 2243
2144 GAAG 2147
2244 GAAG 2247

RESULT 5

US-09-880-107-3338
; Sequence 3338, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3338
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U27460
US-09-880-107-3338

Query Match 78.1%; Score 1676.6; DB 10; Length 1823;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1703; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 225 AGATCTTAGCAAGCAATGTCTCAAGATGGTCTTCTCAGTTTCCAAAGAGTCAATTCGGCA 284
Db 93 AGATCTTAGCAAGCAATGTCTCAAGATGGTCTTCTCAGTTTCCAAAGAGTCAATTCGGCA 152
QY 285 AGAGCTAGAAATTTATCTGTGAAGAGAACTAGAAAAAATACTCACCACAGCATCATCA 344
Db 153 AGAGCTAGAAATTTATCTGTGAAGAGAACTAGAAAAAATACTCACCACAGCATCATCA 212
QY 345 TGAATTTGAGCACCACCAAAAAAGACCTGGATGGATTTCCGAAGCTATTTTCATAGATTTT 404
Db 213 TGAATTTGAGCACCACCAAAAAAGACCTGGATGGATTTCCGAAGCTATTTTCATAGATTTT 272
QY 405 GCAAGAAAGGGGCTCTCTGTGGATTGGGAAAAAATCCAGAGACCCCTCGAAGATTCGAT 464
Db 273 GCAAGAAAGGGGCTCTCTGTGGATTGGGAAAAAATCCAGAGACCCCTCGAAGATTCGAT 332
QY 465 TCAACCTTATGAAAAAGATAAAGCCAGGGCTTGCTGATATATATCTTCCGTGTGAA 524
Db 333 TCAACCTTATGAAAAAGATAAAGCCAGGGCTTGCTGATATATATCTTCCGTGTGAA 392
QY 525 CAACTAGTGGTGGAACTCAATGGTGGTTTGGGAAACCAGCATGGCTGCAAGAGCCC 584
Db 393 CAACTAGTGGTGGAACTCAATGGTGGTTTGGGAAACCAGCATGGCTGCAAGAGCCC 452
QY 585 TAAAGTCTGATGTGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTCAGCAAT 644
Db 453 TAAAGTCTGATGTGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTCAGCAAT 512
QY 645 TGAACATTTGAATAAAACCTCAATACAGATGTTCTCTTGTGTTTAAATGAATCTTTTAA 704
Db 513 TGAACATTTGAATAAAACCTCAATACAGATGTTCTCTTGTGTTTAAATGAATCTTTTAA 572
QY 705 CAGGATGAGATACCAAAAAAATACTACAGAGTACAATCTTCTGTCGTAATAATCTA 764
Db 573 CAGGATGAGATACCAAAAAAATACTACAGAGTACAATCTTCTGTCGTAATAATCTA 632
QY 765 CACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTTACTTCTGTAGCAAA 824
Db 633 CACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTTACTTCTGTAGCAAA 692
QY 825 GGACGTGTTTACTCAGGGGAAAAATACAGAACTGGTACCTCCAGGTCATGGTGATAT 884
Db 693 GGACGTGTTTACTCAGGGGAAAAATACAGAACTGGTACCTCCAGGTCATGGTGATAT 752
QY 885 TTACGCCAGTTTCTACAACTCTGGATTCCTTCATACCTTTATAGAGAGGCAAGAGTA 944
Db 753 TTACGCCAGTTTCTACAACTCTGGATTCCTTCATACCTTTATAGAGAGGCAAGAGTA 812
QY 945 TATTTTGTGCTAAACATAGATAAATCTGGGTGCCAGTGGATCTGTATATTTCTTAATCA 1004
Db 813 TATTTTGTGCTAAACATAGATAAATCTGGGTGCCAGTGGATCTGTATATTTCTTAATCA 872
QY 1005 TCTAATGAACCCACCAATGGAAGAGCGTGTGAATTTGTCTGGAAGTCAAAATAAAC 1064
|||||

Db 873 TCTAATCAACCCACCACCAATGGAAACGCTGTGAATTTCTCATGGAGTCACAATAAAAC 932
Qy 1065 ACGTCAGATGTAAGGGGGGACACTCACTCAATATGAAGCAAACTGAGACTGGTGA 1124
Db 933 ACGTCAGATGTAAGGGGGGACACTCACTCAATATGAAGCAAACTGAGACTGGTGA 992
Qy 1125 AATTGCTCAAGTGCCAAAGACACATGTATAGACAGTTCAAGTCTGTATCAAAAGTTCAAAAT 1184
Db 993 AATTGCTCAAGTGCCAAAGACACATGTGTGACGAGTTCAAGTCTGTATCAAAAGTTCAAAAT 1052
Qy 1185 ATTTAATCAACCAACCTATGATTTCTTTCGACAGCTTAAAGACATGACAGGACAAA 1244
Db 1053 ATTTAATCAACCAACCTATGATTTCTTTCGACAGCTTAAAGACATGACAGGACAAA 1112
Qy 1245 TGCCATTCACATGGAATCATTTGCAATGCAAGACTTTGGATGAGGCTGGAATGTAT 1304
Db 1113 TGCCATTCACATGGAATCATTTGCAATGCAAGACTTTGGATGAGGCTGGAATGTAT 1172
Qy 1305 TCAATTAGAACTGCAGTAGGGGCTGCCATCAAAAGTTTGAATTTCTAGGTATTAA 1364
Db 1173 TCAATTAGAACTGCAGTAGGGGCTGCCATCAAAAGTTTGAATTTCTAGGTATTAA 1232
Qy 1365 TGTGCCAAGGAGCGCTTTCTGCTGTCAAAACCAATCAGATCTTCTGCTGGTATGTC 1424
Db 1233 TGTGCCAAGGAGCGCTTTCTGCTGTCAAAACCAATCAGATCTTCTGCTGGTATGTC 1292
Qy 1425 AAACCTCTATAGTCTTAATGCAAGGATCTTGCAATGAGTGAAGCGGAAATTTCTTAC 1484
Db 1293 AAACCTCTATAGTCTTAATGCAAGGATCTTGCAATGAGTGAAGCGGAAATTTCTTAC 1352
Qy 1485 AGTGCCTTGTGTTAAATAGGAGTCTTTTACGAAGGTTCAAGATTTCTTAAGAAGATT 1544
Db 1353 AGTGCCTTGTGTTAAATAGGAGTCTTTTACGAAGGTTCAAGATTTCTTAAGAAGATT 1412
Qy 1545 TGAAGATATACAGATATGCTTGAATGGATCACTCAGATTTTCAAGGATGTCATTT 1604
Db 1413 TGAAGATATACAGATATGCTTGAATGGATCACTCAGATTTTCAAGGATGTCATTT 1472
Qy 1605 TGAAGAAATGTTTCATTAAGGGAACGGTTATCATTTGCAATCATGTTGACAGAA 1664
Db 1473 TGAAGAAATGTTTCATTAAGGGAACGGTTATCATTTGCAATCATGTTGACAGAA 1532
Qy 1665 TGATATCCCACTGGAGCAGTATTAGAGAAAGATTTGCTGCTGGAACCTTCGATCTT 1724
Db 1533 TGATATCCCACTGGAGCAGTATTAGAGAAAGATTTGCTGCTGGAACCTTCGATCTT 1592
Qy 1725 GGACCACTGAATGAANAATCTGTGGACACTTAAATATGGCTAGTTTCTTACAATGA 1784
Db 1593 GGACCACTGAATGAANAATCTGTGGACACTTAAATATGGCTAGTTTCTTACAATGA 1652
Qy 1785 AATGTTCTTAGGATTTCTAAAATAGGAGTACTTTTACTATGTTACTGTACCTGCAATG 1844
Db 1653 AATGTTCTTAGGATTTAGGCACTAAAGTACTTTTACTATGTTACTGTACCTGCAATG 1712
Qy 1845 TTGATTTTAAATAGATTTTCTGCAATGCTTTTACTGTAAAGAAAGCAGATGGA 1904
Db 1713 TTGATTTTAAATAGATTTTCTGCAATGCTTTTACTGTAAAGAAAGCAGATGGA 1772
Qy 1905 GCAATACTTTCTTCTTGAAGAGATCCCAAGTTAGTTCTTCTTAA 1953
Db 1773 GCAATACTTTCTTCTTGAAGAGATCCCAAGTTAGTTCTTCTTAA 1819

RESULT 6

US-10-102-524-1334
; Sequence 1334, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1334
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-1334

Query Match 26.0%; Score 558.8; DB 9; Length 566;
Best Local Similarity 99.6%; Pred. No. 4.2e-139;
Matches 560; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1293 CCTGAATGTCATTCAATTAGAACTGCAGTAGGGGCTGCCATCAAAAGTTTGAAGATTC 1352
Db 1 CCTGAATGTCATTCAATTAGAACTGCAGTAGGGGCTGCCATCAAAAGTTTGAAGATTC 60
Qy 1353 TCTAGGTATTAAATGTGCCAAGGAGCGCTTTTCTGCTGTCAAAACCAATCAGATCTCTT 1412
Db 61 TCTAGGTATTAAATGTGCCAAGGAGCGCTTTTCTGCTGTCAAAACCAATCAGATCTCTT 120
Qy 1413 GCTGGTATGTCAAACCTCTATAGTCTTAAATGCAAGGATCTCTGCAATGAGTGAAGCG 1472
Db 121 GCTGGTATGTCAAACCTCTATAGTCTTAAATGCAAGGATCTCTGCAATGAGTGAAGCG 180
Qy 1473 GGAATTTCTCAGTGCCTTTGGTTAAATTAGGCAGTTCTTTTACGAAGTTCAAGATTA 1532
Db 181 GGAATTTCTCAGTGCCTTTGGTTAAATTAGGCAGTTCTTTTACGAAGTTCAAGATTA 240
Qy 1533 TCTAAGAAGATTGAAAGTATACCAAGTATGCTTGAATTTGGATCACCTCAGATTTTCAGG 1592
Db 241 TCTAAGAAGATTGAAAGTATACCAAGTATGCTTGAATTTGGATCACCTCAGATTTTCAGG 300
Qy 1593 AGATGTGATTTGGAAAAAATGTTTCAATTAAGGGAACGGTTATCATCATTTGCAAAATCA 1652
Db 301 AGATGTGATTTGGAAAAAATGTTTCAATTAAGGGAACGGTTATCATCATTTGCAAAATCA 360
Qy 1653 TGGTGACAGAAATGATATCCCACTGGAGCAGTATTAGAGAACAGATTTGCTCTGGAAG 1712
Db 361 TGGTGACAGAAATGATATCCCACTGGAGCAGTATTAGAGAACAGATTTGCTCTGGAAG 420
Qy 1713 CCTTCGATCTTGGACCACTGAAATGAAAAATACCTGTGGACACTTAAATAATGGGCTAGT 1772
Db 421 CCTTCGATCTTGGACCACTGAAATGAAAAATACCTGTGGACACTTAAATAATGGGCTAGT 480
Qy 1773 TTCTTACAATGAATTTCTTCTAGGATTTCTAAAATAGGAGGATCTTTTACTATGTTACTG 1832
Db 481 TTCTTACAATGAATTTCTTCTAGGATTTCTAAAATAGGAGGATCTTTTACTATGTTACTG 540
Qy 1833 TACCTCGAGTGTGATTTTAA 1854
Db 541 TACCTCGAGTGTGATTTTAA 562

RESULT 7

US-09-918-995-14100
; Sequence 14100, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0

```
; SEQ ID NO 14100
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; FEATURE:
; LOCATION: (1)...(437)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14100

Query Match      20.2%; Score 432.8; DB 9; Length 437;
Best Local Similarity 99.5%; Pred. No. 1.8e-105;
Matches 434; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1606 GGAAGAAATGTTTCATTAAGGAACGGTTATCATCATTTGCAATTCGAAATCATGGTGCACGAATTT 1665
DB 1 GGAAGAAATGTTTCATTAAGGAACGGTTATCATCATTTGCAATTCGAAATCATGGTGCACGAATTT 60

QY 1666 GATATCCCACTGGAGCAGTATTAGAGAACAGATTGTCTGGAACCTTCGCAATCTTG 1725
DB 61 GATATCCCACTGGAGCAGTATTAGAGAACAGATTGTCTGGAACCTTCGCAATCTTG 120

QY 1726 GACCACTGAATGAAATAATCTGTGGACACTTAAATAATGGGCTAGTTTCTTACAAATGAA 1785
DB 121 GACCACTGAATGAAATAATCTGTGGACACTTAAATAATGGGCTAGTTTCTTACAAATGAA 180

QY 1786 ATGTTCTCTAGGATCTAAATAAGGCGAGTACTTTACTATGTTACTGTACCTCGCAGTGT 1845
DB 181 ATGTTCTCTAGGATCTAAATAAGGCGAGTACTTTACTATGTTACTGTACCTCGCAGTGT 240

QY 1846 TGAATTTTAAATAGAGTTTCTGCAATGATGCTTTTGTAGTCTAAGAAAGCACAGATGGAG 1905
DB 241 TGAATTTTAAATAGAGTTTCTGCAATGATGCTTTTGTAGTCTAAGAAAGCACAGATGGAG 300

QY 1906 CAATACCTTCTCTTTTGAAGAGAAATCCCAAAAGTTAGTTCATCTTAAAGTCAATATTG 1965
DB 301 CAATACCTTCTCTTTTGAAGAGAAATCCCAAAAGTTAGTTCATCTTAAAGTCAATATTG 360

QY 1966 TTTAATCTTAAACTGGGCACTTTGGAAGAACTTTTACAGAACTTTTACAGAGCTCAATGATCA 2025
DB 361 GTTAATCTTAAACTGGGCACTTTGGAAGAACTTTTACAGAACTTTTACAGAGCTCAATGATCA 420

QY 2026 CTTTGAATGCTTGTG 2041
DB 421 CTTTGAATGCTTGTG 436

RESULT 8
US-09-918-995-3419
; Sequence 3419, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3419
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (1)...(430)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-3419

Query Match      20.0%; Score 429; DB 9; Length 430;
Best Local Similarity 99.3%; Pred. No. 6.9e-103;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1105 GCAAACTGAGACTGGTGGAAATTCCTCAAGTGCCAAAGCACATGTAGACGAGTTCAAG 1164
DB 44 GCAAACTGAGACTGGTGGAAATTCCTCAAGTGCCAAAGCACATGTAGACGAGTTCAAG 103

QY 1165 TCTGTATCAAGTTTCAAAATATTTTAAACAAACCTATGATTTCTTTCAGCAGTT 1224
DB 104 TCTGTATCAAGTTTCAAAATATTTTAAACAAACCTATGATTTCTTTCAGCAGTT 163
```

Qy 1225 AAAAGACTGCAGGACCAAAATGCCATTGACATGGAATCATTTGTGAATGCCAAAGACTTTG 1284
Db 164 AAAAGACTGCAGGACCAAAATGCCATTGACATGGAATCATTTGTGAATGCCAAAGACTTTG 223
Qy 1285 GATGAGGCCCTGAATGTCTCAATTTAGAACTGCAGTAGGGCTGCCATCAAAAAGTTTT 1344
Db 224 GATGAGGCCCTGAATGTCTCAATTTAGAACTGCAGTAGGGCTGCCATCAAAAAGTTTT 283
Qy 1345 GAGAAATCTCTAGGTATTAAATGTGCCAAGGAGCCGTTTTCTGCCCTGTCAAAAACCATCA 1404
Db 284 GAGAAATCTCTAGGTATTAAATGTGCCAAGGAGCCGTTTTCTGCCCTGTCAAAAACCATCA 343
Qy 1405 GATCTCTTCTGCTGATGATCAAACTCTATAGTCTTAATGCAAGGATCTTGCAATGAGT 1464
Db 344 GATCTCTTCTGCTGATGATCAAACTCTATAGTCTTAATGCAAGGATCTTGCAATGAGT 403
Qy 1465 GAAAAGCGGGAATTTCTTACAGTGCCTTGGTTAAATAGGCAGTTCTTTTACGAAGTT 1524
Db 404 GAAAAGCGGGAATTTCTTACAGTGCCTTGGTTAAATAGGCAGTTCTTTTACGAAGTT 463
Qy 1525 CAAGATTA 1532
Db 464 CAAGATTA 471

RESULT 10

US-09-918-995-17154
; Sequence 17154, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17154
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-918-995-17154

Query Match 19.7%; Score 422; DB 9; Length 422;
Best Local Similarity 100.0%; Pred. No. 1.3e-102; Mismatches 0; Indels 0; Gaps 0;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1033 TGTGAATTTGTCATGGAAGTCAAAATAAACACGTCGAGATGTAAAGGGCGGGACATC 1092
Db 1 TGTGAATTTGTCATGGAAGTCAAAATAAACACGTCGAGATGTAAAGGGCGGGACATC 60
Qy 1093 ACTCAATATGAAGGCAAACTGAGACTGTGGAAATTTGCTCAAGTCCCAAAAGCAATGTA 1152
Db 61 ACTCAATATGAAGGCAAACTGAGACTGTGGAAATTTGCTCAAGTCCCAAAAGCAATGTA 120
Qy 1153 GACGAGTCAAGTCTGTATCAAAAGTTCAAAATATTATACAAACCACTATGGATTTCT 1212
Db 121 GACGAGTCAAGTCTGTATCAAAAGTTCAAAATATTATACAAACCACTATGGATTTCT 180
Qy 1213 CTTGACGAGTTAAAGACTGCAGGAGCAAAATGCCATTGACATGGAATCATTTGTAAT 1272
Db 181 CTTGACGAGTTAAAGACTGCAGGAGCAAAATGCCATTGACATGGAATCATTTGTAAT 240
Qy 1273 GCAAGACTTTGGATGGAGCCCTGAATGTCAATTTAGAACTGCAGTAGGGCTGCC 1332
Db 241 GCAAGACTTTGGATGGAGCCCTGAATGTCAATTTAGAACTGCAGTAGGGCTGCC 300
Qy 1333 ATCAAAAGTTTTGAGAAATTTCTAGGTATTATGTGCCAAGGAGCGGTTTTCTTGCCTGTC 1392
Db 301 ATCAAAAGTTTTGAGAAATTTCTAGGTATTATGTGCCAAGGAGCGGTTTTCTTGCCTGTC 360

Qy 1393 AAAACCAATCAGATCTCTTCTGCTGATGTCAAACTCTATAGTCTTTAATGCAGATCT 1452
Db 361 AAAACCAATCAGATCTCTTCTGCTGATGTCAAACTCTATAGTCTTTAATGCAGATCT 420
Qy 1453 CT 1454
Db 421 CT 422

RESULT 11

US-09-918-995-36518
; Sequence 36518, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36518
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-918-995-36518

Query Match 19.6%; Score 421.4; DB 9; Length 433;
Best Local Similarity 99.8%; Pred. No. 2e-102; Mismatches 1; Indels 0; Gaps 0;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 30 GTCGGAGCCGCTGACGGGTGGACAGGGGGGTTAGCAGCTGGGCTGCGACCCGTTAGG 89
Db 11 GCGGGAGCCGCTGACGGGTGGACAGGGGGGTTAGCAGCTGGGCTGCGACCCGTTAGG 70
Qy 90 GAGGGCTCAAGGTGTGCATGTGTGAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 149
Db 71 GAGGGCTCAAGGTGTGCATGTGTGAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 130
Qy 150 TGACTTTTCAGCTGCGAGCTTTCTTCCGGGGCGCCATAAACCGCCCCCAATTTCCAGCT 209
Db 131 TGACTTTTCAGCTGCGAGCTTTCTTCCGGGGCGCCATAAACCGCCCCCAATTTCCAGCT 190
Qy 210 GCTAAGGAAGAGAGAGATCTTAGCAAGCAATGTCTCAAGATGCTTCTCAGTTCCA 269
Db 191 GCTAAGGAAGAGAGAGATCTTAGCAAGCAATGTCTCAAGATGCTTCTCAGTTCCA 250
Qy 270 AGAAGTCAATTCGGCAAGAGCTAGAAATTTCTGTGAAGAGGAACCTAGAAAAAATCTAC 329
Db 251 AGAAGTCAATTCGGCAAGAGCTAGAAATTTCTGTGAAGAGGAACCTAGAAAAAATCTAC 310
Qy 330 CACAGCATCATCATGAATTTTGAGCACACCAAAAAAGACCTTGGATGGAATTTTCGAAGCT 389
Db 311 CACAGCATCATCATGAATTTTGAGCACACCAAAAAAGACCTTGGATGGAATTTTCGAAGCT 370
Qy 390 ATTTCATAGATTTTTCGAAGAAAAAGGGGCTTTCTGTGATTTGGGAAAAAATCCAGAGACC 449
Db 371 ATTTCATAGATTTTTCGAAGAAAAAGGGGCTTTCTGTGATTTGGGAAAAAATCCAGAGACC 430
Qy 450 CCC 452
Db 431 CCC 433

RESULT 12

US-09-917-800A-877/c
; Sequence 877, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark
 ; APPLICANT: Johnson, Kory
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5038-US
 ; CURRENT APPLICATION NUMBER: US 09/917,800A
 ; CURRENT FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,040
 ; PRIOR FILING DATE: 2000-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,880
 ; PRIOR FILING DATE: 2000-11-02
 ; PRIOR APPLICATION NUMBER: US 60/290,029
 ; PRIOR FILING DATE: 2001-05-11
 ; PRIOR APPLICATION NUMBER: US 60/290,645
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: US 60/292,336
 ; PRIOR FILING DATE: 2001-05-22
 ; PRIOR APPLICATION NUMBER: US 60/295,798
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 60/297,457
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,884
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: US 60/303,459
 ; PRIOR FILING DATE: 2001-07-09
 ; NUMBER OF SEQ ID NOS: 1740
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 877
 ; LENGTH: 671
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 A1170679
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(671)
 ; OTHER INFORMATION: n = a or c or g or t
 ; US-09-917-800A-877

Query Match 19.6%; Score 419.8; DB 10; Length 671;
 Best Local Similarity 81.3%; Pred. No. 7e-102;
 Matches 536; Conservative 0; Mismatches 114; Indels 9; Gaps 4;
 QY 1403 CAGATCTTGTGTTGATGTCCTAACTCTATAGTCTTAATGACGAGATCTTGACATGA 1462
 DB 671 CAGATCTTGTGTTGATGTCCTAACTCTTAACGCGGATCTTTGACCATGA 612
 QY 1463 GTGAAAGCGGAATTTCTACAGTCGCTTGGTTAAATAGGACGTTCTTTTACGAAG 1522
 DB 611 GTGAAAGCGGTAATTTCTACAGTACCTTGGTTAAATAGGACGTTCTTTTACGAAG 552
 QY 1523 TTCAAGATTATCTAAGAAATTTGAAAGTATACAGATATGCTTGAATTTGGATCACTCA 1582
 DB 551 TTCAAGATTATCTAAGAAATTTGAAAGTATACCCGATATGTTGAACTGGACCACTCA 492
 QY 1583 CAGTTTCAGGAGATGTGACATTTGGAAAAATGTTTCAATTAAGGAAACGGTTATCATCA 1642
 DB 491 CGGTTTCAGGAGATGTAACTTTGGAAAAATGTTTCAATTAAGGAAACAGTTATCATCA 432
 QY 1643 TTGCAATCATGTGTGACAGATTTGATATCCACTGGACGATTTAGAGAACAGATTG 1702
 DB 431 TTGCAATCATGTGTGACAGATTTGATATCCCTCCCGAGGACGATTTAGAGAACAGATTG 372
 QY 1703 TGTCTGAAACCTTCGACATCTTGACACCTGAAATGAAAAATCTGTGGACACTT----- 1757
 DB 371 TATCTGGAACTTCGACATCTTGACACCTGAAATGAGCACTGTGACCTACACTTTCTAC 312
 QY 1758 AATAATGGCTAGTTTCTTACAAATGAATGTTCTCTAGGATTTCTAAATAGGACGGTAC 1817
 DB 311 TAATATGGCTAAAGATTTTACAAATGAATGTTCTCTAGGATTTCTAAATAGGACGGTAC 252

QY 1818 TTTA-CTATGTTACTGTACCTCGAGTGTGATTTTAAAAATAGAGTTTTCTGCAGTATG 1876
 DB 251 TTTATTTACTATGTTGTTACTCTGCACATTTGATTTTAAAGTA--GTTTTCTGCAATGAG 194
 QY 1877 CTTTGTAGTCTAAGAAAGCAGATGAGGAAATCTTTCTTCTTTGAAGAGATCCCAA 1936
 DB 193 CTTTGTAGTCTAAGAAAGCCTTAAACAAAGCAATCTTTCTTCTTTGGATGAAAGCCTAG 134
 QY 1937 AAGTTAGTTCATCTTAAAGTCAATATTGTTTAAATCTTAAACCTGGGCAACTTTGGAAGA 1996
 DB 133 AAATTAGTCTCTTTAAAGTCAATATTATTATCTCAGAAATCGGGCCAGCTCTGTAGA 74
 QY 1997 ACTTTTAAAGAGCCTCAATGATGATCACTTTGAAATGCTTTGTTGATTTCAAAAAATAA 2055
 DB 73 TC-TTCAACAGAGCTTCAGTGACTGCTCCCTTTGAATTCCTTTGATTTCAAAAAATAA 16
 RESULT 13
 US-09-918-995-8128
 ; Sequence 8128, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8128
 ; LENGTH: 439
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(439)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-918-995-8128

Query Match 18.1%; Score 388; DB 9; Length 439;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1024 GGAAACCGTGTGAATTTGTCATGAAGTCAAAATAAAACGTCGAGATGTAAGGCG 1083
 DB 52 GGAAACCGTGTGAATTTGTCATGAAGTCAAAATAAAACGTCGAGATGTAAGGCG 111
 QY 1084 GGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGAATTTGCTCAAGTGCCAAA 1143
 DB 112 GGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGAATTTGCTCAAGTGCCAAA 171
 QY 1144 GCACATGTAGACAGTTCAAGTCTGTATCAAAATTTCAAAATATTTAATACAAAACCTTA 1203
 DB 172 GCACATGTAGACAGTTCAAGTCTGTATCAAAATTTCAAAATATTTAATACAAAACCTTA 231
 QY 1204 TGGATTTCTTTGACAGATTTAAAGAGCTGAGGAGCAAAATGCCATTTGACATGGAATC 1263
 DB 232 TGGATTTCTTTGACAGATTTAAAGAGCTGAGGAGCAAAATGCCATTTGACATGGAATC 291
 QY 1264 ATTGTGAATGCAAGACTTTGATCGGCGCTGAATGTCATTTCAATTTAGAAATCGCAGTA 1323
 DB 292 ATTGTGAATGCAAGACTTTGATCGGCGCTGAATGTCATTTCAATTTAGAAATCGCAGTA 351
 QY 1324 GGGGCTGCCATCAAAAGTTTTCAGAAATTTCTAGGTATTAAATGTGCAAGAGCGGTTTT 1383
 DB 352 GGGGCTGCCATCAAAAGTTTTCAGAAATTTCTAGGTATTAAATGTGCAAGAGCGGTTTT 411
 QY 1384 CTGCTGTCAAAACACATCAGATCTCT 1411
 DB 412 CTGCTGTCAAAACACATCAGATCTCT 439

RESULT 14

US-09-960-352-6825
; Sequence 6825, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6825
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 29-LIB34-050-Q1-E1-H1
US-09-960-352-6825

Query Match 18.0%; Score 386.2; DB 10; Length 428;
Best Local Similarity 94.6%; Pred. No. 5.1e-93;
Matches 400; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY	1180	AAAAATATTATACAAACCACTATGGATTCTCTTGCAGCAGTTAAAGACTGCAGGAG	1239
DB	6	AAAAATATTATACAAACCACTATGGATTCTCTTGTGCGAGTTAAAGACTGCAGGAG	65
QY	1240	CAAAATGCCATTGACATGGAATCATTGTGAATGCAAGACATTTGGATGGAGGCTGAAT	1299
DB	66	CAGATGCTATTGACATGGAATCATTGTGAATGCAAGACATTTGGATGGAGGCTGAAT	125
QY	1300	GTCATTCAATTAGAACTGCAGTAGGGCTGCCATCAAAAGTTTGGAGATTTCTTAGGT	1359
DB	126	GTTATTTCAGTTAGAACTGCAGTAGGGCTGCCATCAAAAGTTTGGAGATTTCTTAGGT	185
QY	1360	ATTATGTGCCAAGGAGCGTTTCTGCTGTCAAAACCAATCAGATCTCTTGTGTTG	1419
DB	186	ATTATGTTCCTAGAACGCGTTTCTGCTGTCAAAACCAATCAGATCTCTTGTGTTG	245
QY	1420	ATGTCAAACCTCTATAGTCTTAATGCAGGATCTGTGACAAATCAGTGAAGCGGAATTT	1479
DB	246	ATGTCAAACCTCTACAGCTTAATGCAGGATCTGTGACAAATCAGTGAAGCGGAATTT	305
QY	1480	CCTACAGTCCCTTGGTTAAATTAGCAGTCTTTTACGAAGTTCAAGATTATCTAAGA	1539
DB	306	CCTACAGTCCCTTGGTTAAATTAGCAGTCTTTTACGAAGTTCAAGATTATCTAAGA	365
QY	1540	AGATTGAAAGTATACCAATATGCTTGAATGGATCACTCAGTTTCAGAGATGTG	1599
DB	366	AGATTGAAAGTATACCAATATGCTTGAATGGATCACTCAGTTTCAGAGATGTG	425
QY	1600	ACA 1602	
DB	426	ACA 428	

RESULT 15

US-09-970-966-43/c
; Sequence 43, Application US/09970966
; Patent No. US20020173638A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.484C6

; CURRENT APPLICATION NUMBER: US/09/970.966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 108_195, 213, 279, 287, 349
; OTHER INFORMATION: n = A,T,C or G
US-09-970-966-43

Query Match 17.5%; Score 375.2; DB 9; Length 396;
Best Local Similarity 97.7%; Pred. No. 4.2e-90;
Matches 377; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY	1684	GTATTAGAGAACAGATTGTCTGGAAACCTTGGACCATCTGGACCACTGAATGAAAAA	1743
DB	396	GTATTAGAGAACAGATTGTCTGGAAACCTTGGACCATCTGGACCACTGAATGAAAAA	337
QY	1744	TACTGTGGACACTTAAATAATGGGCTAGTTTCTTACAAATGAAATGTTCTTAGGATTCTA	1803
DB	336	TACTGTGGACACTTAAATAATGGGCTAGTTTCTTACAAATGAAATGTTCTTAGGATTCTA	277
QY	1804	AAATAGGCAGGTACTTTTACTATGTGTACCTGCGAGTGTGATTTTTTAAATAGAGT	1863
DB	276	AAATAGGCAGGTACTTTTACTATGTGTACCTGCGAGTGTGATTTTTTAAATAGAGT	217
QY	1864	TTTCTGCGAGTATGCTTTTAGTCTTAAGAAAGCAGATGGAGCAATACTTTTCTTTTG	1923
DB	216	TTTNTGCGAGTATGCTTTTAGTCTTAAGAAAGCAGATGGAGCAATACTTTTCTTTTG	157
QY	1924	AAGCAATCCCAAGTTAGTTTCATCTTAAAGTCAATATTGTTTAACTTAAACTGGG	1983
DB	156	AAGCAATCCCAAGTTAGTTTCATCTTAAAGTCAATATTGTTTAACTTAAACTGGG	97
QY	1984	CAACTTTTGGAAAGAACTTTTAAACAGAGCCTCAATGATGATCACTTTGATTTGAT	2043
DB	96	CAACTTTTGGAAAGAACTTTTAAACAGAGCCTCAATGATGATCACTTTGATTTGAT	37
QY	2044	TTCAAAAATAAAGCAGTGAAGCAATA 2069	
DB	36	TTCAAAAATAAAGCAGTGAAGCAATA 11	

Search completed: June 28, 2003, 12:28:51
Job time : 206.744 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:12:40 ; Search time 1985.1 Seconds

(without alignments)
17516.326 Million cell updates/sec

Title: US-09-981-353-104

Perfect score: 2147

Sequence: 1 gtgttggagaaagtaggg.....gttactatattgttgaag 2147

Scoring table: IDENTITY_NUC

Gapop:10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estrov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1830.8	85.3	2128	11	BC000173 Homo sapi
2	972	45.3	1115	13	BM473685 AGENCOURT
3	936	43.6	959	9	AL533518 AL533518
4	935.6	43.6	973	9	AL513738 AL513738
5	929.6	43.3	945	9	AL557910 AL557910
6	915.8	42.7	979	9	AL579851 AL579851

7	914.6	42.6	1182	13	BM473578
8	878.4	40.9	926	9	AL541422
9	875	40.8	965	9	AL561196
10	873.2	40.7	1012	14	BQ072394
11	850.8	39.6	908	14	BQ213239
12	841.8	39.2	857	9	AL568945
13	810	37.1	872	14	BQ223934
14	797.4	37.1	823	9	AL565147
15	791.6	36.9	845	9	AU119264
16	791	36.8	861	12	BG754488
17	789.6	36.8	930	9	AL561167
18	786	36.6	842	9	AU132957
19	784.4	36.5	868	9	AU121194
20	778.2	36.2	811	10	BE617619
21	765	35.6	792	9	AL541421
22	761.2	35.5	790	13	BM014590
23	755	35.2	805	12	BG716738
24	753.8	35.1	849	14	BQ227374
25	750.4	35.0	763	13	BI333577
26	747.4	34.8	855	9	AU122795
27	744.2	34.7	905	14	BQ431692
28	740	34.5	911	12	BG566133
29	737.8	34.4	776	13	BI093754
30	737	34.3	747	9	AU134523
31	730.4	34.0	919	13	BI601687
32	729.8	34.0	867	9	AL559311
33	729.2	34.0	833	12	BG570842
34	724.6	33.7	1147	14	BM799747
35	721	33.6	738	13	BI547723
36	721	33.6	976	10	BE617700
37	717.6	33.4	873	9	AL580952
38	714.4	33.3	723	9	AU126119
39	713.6	33.2	831	9	AI198572
40	707.6	33.0	728	14	BQ027133
41	706.4	32.9	748	9	AU130360
42	704.2	32.8	772	12	BG428491
43	698.4	32.5	700	9	AA773559
44	698.4	32.5	714	14	BM997100
45	696.8	32.5	1118	13	BM543980

ALIGNMENTS

RESULT 1
BC000173
LOCUS BC000173 2128 bp mRNA linear HTC 12-JUL-2001
DEFINITION Homo sapiens, UDP-glucose pyrophosphorylase 2, clone IMAGE:2985723,
mRNA.
ACCESSION BC000173 GI:12652840
VERSION BC000173.1
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2128)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Db 2008 ATTGTTTAACTTAAACCTGGCAACTTTTGAAGAACTTTTAAACAGAGCCTCAATGATG 2067

Qy 2022 ATCACTTTGAATTCCTGATTTTCAAAAATAAAGCAGTGAAGCAATA 2069

Db 2068 ATCACTTTGAATTCCTGATTTTCAAAAATAAAGCAGTGAAGCAATA 2115

RESULT 2

BM473685

LOCUS

DEFINITION AGENCOURT_6456577 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5562065

5', mRNA sequence.

BM473685

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1115)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM12290 row: n column: 18

High quality sequence stop: 733.

Location/Qualifiers

1..1115

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5562065"

/clone.lib="NIH MGC 88"

/tissue_type="duodenal adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: Small intestine; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."

BASE COUNT 364 a 205 c 234 g 310 t 2 others

ORIGIN

Query Match 45.3%; Score 972; DB 13; Length 1115;

Best Local Similarity 96.1%; Pred. No. 4.8e-255;

Matches 1070; Conservative 0; Mismatches 37; Indels 7; Gaps 7;

Qy 632 CTGTTACAGCAATTTGAATTTAAATTAACCTTACAAATACAGATGTTCTCTGTTTAA 691

Db 1 CTGTTACAGCAATTTGAATTTAAATTAACCTTACAAATACAGATGTTCTCTGTTTAA 60

Qy 692 TGAACCTTTTAAACGGGATGAGATACCAAAAATATCTACAGATCAATCATTTGTC 751

Db 61 TGAACCTTTTAAACGGGATGAGATACCAAAAATATCTACAGATCAATCATTTGTC 120

Qy 752 GTGTGAAAATCTACACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTAC 811

Db 121 GTGTGAAAATCTACACTTTCAATCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTAC 180

Qy 812 TTCTCTAGCAAGGACGTGTTCTTACTCAGGGGAAAATACAGAGCTTGGTACCTCCAG 871

Db 181 TTCTCTAGCAAGGACGTGTTCTTACTCAGGGGAAAATACAGAGCTTGGTACCTCCAG 240

Qy 872 GTCATGGTATATTACGCCAGTTTCTACAACTCTGGATTGCTTGATCTTTATAGGAG 931

Db 241 GTCATGGTATATTACGCCAGTTTCTACAACTCTGGATTGCTTGATCTTTATAGGAG 300

Qy 932 AAGCAAAGAGTATATTTTGTGTCTAACATAGATAAATCTGGTGCCACAGTGGATCTGT 991

Db 301 AAGCAAAGAGTATATTTTGTGTCTAACATAGATAAATCTGGTGCCACAGTGGATCTGT 360

Qy 992 ATATTTCTTAATCATCTAATGAACCCCAATGAAACCGCTGTGAATTTCTCATGGAG 1051

Db 361 ATATTTCTTAATCATCTAATGAACCCCAATGAAACCGCTGTGAATTTCTCATGGAG 420

Qy 1052 TCACAAATAAAACACGCTGCAGATGTAAGGGCGGGACACTCCTCAATATGAAGCAAAAC 1111

Db 421 TCACAAATAAAACACGCTGCAGATGTAAGGGCGGGACACTCCTCAATATGAAGCAAAAC 480

Qy 1112 TGACACTGGTGGAAATTTGCTCAAGTGCACCAAGCACATGTAGACAGTTCAAGTCTGTAT 1171

Db 481 TGACACTGGTGGAAATTTGCTCAAGTGCACCAAGCACATGTAGACAGTTCAAGTCTGTAT 540

Qy 1172 CAAAGTTCCAAATATTTAATACAAACCACTTATGGATTTCTCTTGCAGCAGTTAAAGAC 1231

Db 541 CAAAGTTCCAAATATTTAATACAAACCACTTATGGATTTCTCTTGCAGCAGTTAAAGAC 600

Qy 1232 TGCAGGAGCAAAATGCCATTGACATGGAATCATTTGGAATGCAAGAACTTTGATGGAG 1291

Db 601 TGCAGGAGCAAAATGCCATTGACATGGAATCATTTGGAATGCAAGAACTTTGATGGAG 660

Qy 1292 GCCTGAATGTCAATTTCAATTTAGAAAATCGAGTAGGGGCTGCCATCAAAAAGTTTGAAGATT 1351

Db 661 GCCTGAATGTCAATTTCAATTTAGAAAATCGAGTAGGGGCTGCCATCAAAAAGTTTGAAGATT 720

Qy 1352 CTCTAGGTATTAATGTGCCAAGGAGCCGTTTTCCTCTCAAAACCACTCAGATCTCT 1411

Db 721 CTCTAGGTATTAATGTGCCAAGGAGCCGTTTTCCTCTCAAAACCACTCAGATCTCT 779

Qy 1412 TGCTGGTGTATGTCACCAACCTCTATAGTCTTAATGAGGATCTCTGACAAATAGTGAAAAGC 1471

Db 780 TGCTGGTGTATGTCACCAACCTCTATAGTCTTAATGAGGATCTCTGACAAATAGTGAAAAGC 839

Qy 1472 GGGAAATTTCTTACAGTGCCCTTGGTTAAATTAGGCAG-TTCTTTTACGAAGTTTCAAGAT 1530

Db 840 GGGAAATTTCTTACAGTGCCCTTGGTTAAATTAGGCAGTTTCTTTTACGAAGGTTTCAAGA 899

Qy 1531 TATCTAGAGATTTGAAAGTATACCATATGCTTGAATTTGGATCCTCCTCAGTTTCA 1590

Db 900 TATCTAGAGATTTGAAAGTATACCATATGCTTGAATTTGGATC-CCTCAGATTTTC 958

Qy 1591 GGAGATGTGACATTTGGAAAAATTTTTCATTAAAGGGAACGGTTTATCATCATTCGAAAT 1650

Db 959 AGGAGATGTACAAATTTGAAAAATTTCCCTTAAAGGAAACGTTTATC-TCATTGCAAT 1017

Qy 1651 CATGTGACAGATTTGATATCCCTCTGGAGCAGTATTAGAGAACAAAGATTGTCTGGA 1710

Db 1018 CAGGGTGAC-GAATTTGATATCCCTCTGGGAGC-GTTTTTGAAGAACAGGATGGTTCTGGG 1075

Qy 1711 AACCTTCATCTTGGACCACTGAAATGAAAAAT 1744

Db 1076 AACCTTC-CCCTTGGACCTTTGAAAGGAAAAAT 1108

RESULT 3

AL533518

LOCUS

DEFINITION AL533518 LTI FL015_Brn1 Homo sapiens cDNA clone CSODN0041M14 5 prime, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 959)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

AL533518

959 bp mRNA linear EST 13-FEB-2001

AL533518 LTI FL015_Brn1 Homo sapiens cDNA clone CSODN0041M14 5 prime, mRNA sequence.

AL533518

GI:12797011

EST.

human.

Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Location/Qualifiers
1..959
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DN004YM14"
/clone_lib="JTI_FL015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies.
Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 309 a 177 c 208 g 263 t 2 others
ORIGIN

Query Match	43.6%	Score 936;	DB 9;	Length 959;
Best Local Similarity	99.5%;	Pred. No. 3.3e-245;	Mismatches 2;	Indels 1;
Matches 947;	Conservative	2;	Gaps	1;

```

Qy      542 AACTCAATGGTGTTGCGGAACCAAGCATGGCTGCAGAAAGGCCCTAAAAGTCTGATTGGTG 601
       |||
Db      9 AAMTCATGGTGTTGTTGGGAA-CAGCATGGCTCAAAGGCCCTAAAAGTCTGATTGGTG 67

Qy     602 TGAGGAATGAGAATACCTTTCTGTGACTGCTGTTCAGCAAAATTGACAATTTGNAATAAAA 661
       |||
Db     68 TGAGGAATGAGAATACCTTTCTGGATCTGACTGTTCAGCAAAATTGAAACATTTGAATAAAA 127

Qy     662 CCTACAATACAGATGTTCCCTCTCTTTTAATGAACTCTTTTTAACACGGATGAAGATACCA 721
       |||
Db    128 CTTACAATACAGATGTTCCCTCTCTTTTAATGAACTCTTTTAACACGGATGAAGATACCA 187

Qy     722 AAAAAAATACACAGAAGTACAATCATTTGTCGTGTGAAAAATCTACATTTTCAATCAAAAGCA 781
       |||
Db    188 AAAAAAATACACAGAAGTACAATCATTTGTCGTGTGAAAAATCTACATTTTCAATCAAAAGCA 247

Qy     782 GGTACCCGAGGATTAATAAAGATCTTTACTTCTGTAGCAAGGACGTGTCTTACTCAG 841
       |||
Db    248 GGTACCCGAGGATTAATAAAGATCTTTTACTTCTCTGTAGCAAGGACGTGTCTTACTCAG 307

Qy     842 GGGAAAAATACAGAAGCTTGGTACCCTCCAGTCTAGTGTATTTTACGCCAGTTTCTTACA 901
       |||
Db    308 GGGAAAAATACAGAAGCTTGGTACCCTCCAGTCTAGTGTATTTTACGCCAGTTTCTTACA 367

Qy     902 ACTCTGGATTGCTTGATACCTTTATAGGAGAAGGCAAGAGTAGTATATTTTGTGTCTTAACA 961
       |||
Db   368 ACTCTGGATTGCTTGATACCTTTATAGGAGAAGGCAAGAGTAGTATATTTTGTGTCTTAACA 427

Qy     962 TAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCACCCA 1021
       |||
Db   428 TAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCACCCA 487

Qy    1022 ATGGAAAACCGTGTGNAATTTGTCTATGGAGTCCACAAATAAACACGTCGAGATGTAAGG 1081
       |||
Db   488 ATGGAAAACCGTGTGNAATTTGTCTATGGAGTCCACAAATAAACACGTCGAGATGTAAGG 547

Qy    1082 CGCGGACACTCACATCAATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCTCA 1141
       |||
Db   548 CGCGGACACTCACATCAATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGTGCCTCA 607

Qy    1142 AAGCACATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAATATTTTAATACAAAACACC 1201
       |||
Db   608 AAGCACATGTAGACGAGTTCAAGTCTGTATCAAAAGTTCAAATATTTTAATACAAAACACC 667

```

JOURNAL COMMENT		Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.	
FEATURES		Location/Qualifiers 1. 945 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CS0DJ003YP10" /clone_lib="LTI_NFL008_TC2" /sex="male" /tissue_type="T cells from T cell leukemia" /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"	
BASE COUNT		294 a 180 c 231 g 239 t 1 others	
ORIGIN			
Query Match 43.3%; Score 929.6; DB 9; Length 945; Best Local Similarity 99.8%; Pred. No. 1.9e-243; Matches 940; Conservative 1; Mismatches 0; Indels 1; Gaps 1;			
Qy	54	CAAGGGGGGTTAGCAGCTGGGCTGGACCGTTAGGGAGGGGCTCAAGTGTGCATGTGT	113
Db	1	CAAGGGGGGTTAGCAGCTGGGCTGGACCGTTAGGGAGGGGCTCAAGTGTGCATGTGT	60
Qy	114	GAGGGAAGAGAGAGAGAGAGAGGGCGCTCAGAGGTGACTTTTCAGCCTCGAGGCTTCT	173
Db	61	GAGGGAAGAGAGAGAGAGAGAGGGCGCTCAGAGGTGACTTTTCAGCCTCGAGGCTTCT	120
Qy	174	TCCCGGGGGCCCATAAACGGCCCCCAATTTCCAGCTGCTAAAGGAAGAGAGATCTTAG	233
Db	121	TCCCGGGGGCCCATAAACGGCCCCCAATTTCCAGCTGCTAAAGGAAGAGAGATCTTAG	180
Qy	234	CAAAAGCAATGCTCTCAAGATGGTGTCTCAGTTCCTCAAGAAAGTCAATTCGGCAAGAGCTAGA	293
Db	181	CAAAAGCAATGCTCTCAAGATGGTGTCTCAGTTCCTCAAGAAAGTCAATTCGGCAAGAGCTAGA	240
Qy	294	ATTATCTGTGAAGAGAACTAGAAAAATACTCACCACAGCATCATCAGATGAATTTGA	353
Db	241	ATTATCTGTGAAGAGAACTAGAAAAATACTCACCACAGCATCATCAGATGAATTTGA	300
Qy	354	GCACACAAAAAGACCTGGATGGATTTCCGAAGCTATTTTCATAGATTTTTCGAAGAAA	413
Db	301	GCACACAAAAAGACCTGGATGGATTTCCGAAGCTATTTTCATAGATTTTTCGAAGAAA	360
Qy	414	GGGGCTTCTGTGGATTTGGGAAAAATCCAGAGACCCCTCGAAGATTCGATTTCAACCTTA	473
Db	361	GGGGCTTCTGTGGATTTGGGAAAAATCCAGAGACCCCTCGAAGATTCGATTTCAACCTTA	420
Qy	474	TGAAAAAGATAAGCCAGGGGCTTGCCTGATATATATCTTCCGTGTTTGACAAACTAGT	533
Db	421	TGAAAAAGATAAGCCAGGGGCTTGCCTGATATATATCTTCCGTGTTTGACAAACTAGT	480
Qy	534	GGTGTGAAACTCAATGGTGTGTTGGGAACCCAGCATGGGCTGCAAGAGCCCTTAAAGTCT	593
Db	481	GGTGTGAAACTCAATGGTGTGTTGGGAACCCAGCATGGGCTGCAAGAGCCCTTAAAGTCT	540
Qy	594	GATTGGTGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGCTTTCAGCAAAATGAACATTT	653
Db	541	GATTGGTGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGCTTTCAGCAAAATGAACATTT	600
Qy	654	GAATAAAACCTTACAAATCAGATGTTCTCTTGTGTTTAAATGAATCTTTTAAACGGATGA	713
Db	601	GAATAAAACCTTACAAATCAGATGTTCTCTTGTGTTTAAATGAATCTTTTAAACGGATGA	660

QY 714 AGATACCAAAATACTACAGAGTACAAATCATTTGTCGTGGAATACTACACTTTCAA 773
 Db 661 AGATACCAAAATACTACAGAGTACAAATCATTTGTCGTGGAATACTACACTTTCAA 720
 QY 774 TCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTACTTCTCTAGCAAGGACGTGTC 833
 Db 721 TCAAGCAGGTACCCGAGGATTAATAAGAAATCTTTACTTCTCTAGCAAGGACGTGTC 780
 QY 834 TTACTCAGGGGAAATACAGAGCTTGGTACCTCAGGTATGTTGATATTTAGCCGAG 893
 Db 781 TTACTCAGGGGAAATACAGAGCTTGGTACCTCAGGTATGTTGATATTTAGCCGAG 840
 QY 894 TTCTTACAACTCGATTGCTTGTATACCTTTATAGGAGGCAAGAGTATATTTTGT 953
 Db 841 TTCTTACAACTCGATTGCTTGTATACCTTTATAGGAGGCAAGAGTATATTTTGT 900
 QY 954 GTCTAACATAGATAATCTGGTGCCACAGTGGATCTGTATAT 995
 Db 901 GTCTWACATAGATAATCTGGTG-CACAGTGGATCTGTATAT 941

RESULT 6
 AL579851/c
 LOCUS AL579851 LTI_NFL008_Tc2 Homo sapiens cDNA clone CSODJ003YP10 3
 DEFINITION prime, mRNA sequence.
 ACCESSION AL579851
 VERSION AL579851.1 GI:12945297
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 979)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1..979
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODJ003YP10"
 /clone_lib="LTI_NFL008_Tc2"
 /sex="male"
 /tissue_type="T cells from T cell leukemia"
 /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 276 a 194 c 176 g 311 t 22 others
 ORIGIN

Query Match 42.7%; Score 915.8; DB 9; Length 979;
 Best Local Similarity 95.6%; Pred. No. 1.2e-239;
 Matches 925; Conservative 22; Mismatches 20; Indels 1; Gaps 1;
 QY 1028 AACGCTGTGAATTTGTCATGGAAGTCACAAATAAACACAGTCAGATGTAAAGGGCGGA 1087
 Db 967 AAACGCTKTTTTCATGGAAGTCACAAATAAACACAGTCAGATGTAAAGGGC-GGA 909
 QY 1088 CACTCACTCAATATGAAGGCAAACTGAGACTGGTGAAATTTGCTCAAGTGCCAAAGCAC 1147

Db 908 CACTCACTCAATATGAAGGCAAACTGAGACTGGTGAAATTTGCTCAAGTGCCAAAGCAC 849
 QY 1148 ATGTAGACGAGTTCAGAGTCTGTATCAAGTTCAAAATATTTAATACAAACACCTATGGA 1207
 Db 848 ATGTAGACGAGTTCAGAGTCTGTATCAAGTTCAAAATATTTAATACAAACACCTATGGA 789
 QY 1208 TTTCTTCTGACAGCTTAAAGAGCTGCAGGAGCAAAATGCCATTGACATGGAATCATTTG 1267
 Db 788 TTTCTTCTGACAGCTTAAAGAGCTGCAGGAGCAAAATGCCATTGACATGGAATCATTTG 729
 QY 1268 TGAATGCAAGAGCTTTGGATGGAGCCCTGAATGTCATTTCAATAGAACTCGAGTAGGG 1327
 Db 728 TGAATGCAAGAGCTTTGGATGGAGCCCTGAATGTCATTTCAATAGAACTCGAGTAGGG 669
 QY 1328 CTGCCATCAAAAGTTTGGATGGAGCTTCTAGTATTTAATGTGCCAGGAGCCGTTTTCTGC 1387
 Db 668 CTGCCATCAAAAGTTTGGATGGAGCTTCTAGTATTTAATGTGCCAGGAGCCGTTTTCTGC 609
 QY 1388 CTGTCAAAACCAACATCAGATCTTCTGTGTGTGATGTCAAACTCTATAGTCTTAAATGCAG 1447
 Db 608 CTGTCAAAACCAACATCAGATCTTCTGTGTGTGATGTCAAACTCTATAGTCTTAAATGCAG 549
 QY 1448 GATCTCTGACAAATGAGTGAAGGCGGAATTTCTACAGTGCCTTGGTTAAATTAGGCA 1507
 Db 548 GATCTCTGACAAATGAGTGAAGGCGGAATTTCTACAGTGCCTTGGTTAAATTAGGCA 489
 QY 1508 GTTCTTTTACCAAGGTTCAAGATTTATCAAGAAGTTTGAAGTATACAGATATGCTTG 1567
 Db 488 GTTCTTTTACCAAGGTTCAAGATTTATCAAGAAGTTTGAAGTATACAGATATGCTTG 429
 QY 1568 AATTGGATCACTCACAGTTTTCAGGAGATGTGACATTTGGAAAAAATGTTTCATTAAAGG 1627
 Db 428 AATTGGATCACTCACAGTTTTCAGGAGATGTGACATTTGGAAAAAATGTTTCATTAAAGG 369
 QY 1628 GAACGGTTATCATTCATTCGCAAAATCATGTGACAGATTTGATATCCACCTCGAGCAGTAT 1687
 Db 368 GAACGGTTATCATTCATTCGCAAAATCATGTGACAGATTTGATATCCACCTCGAGCAGTAT 309
 QY 1688 TAGAACAAGATTTGTCTGGAACCTTCCACCTTGGACCACTGAATGAATAACT 1747
 Db 308 TAGAACAAGATTTGTCTGGAACCTTCCACCTTGGACCACTGAATGAATAACT 249
 QY 1748 GTGGACACTTAAATAATGGCTAGTTTCTTACAATGAATGTTCTAGGATTTCTAAAT 1807
 Db 248 GTGGACACTTAAATAATGGCTAGTTTCTTACAATGAATGTTCTAGGATTTCTAAAT 189
 QY 1808 AGGAGGTACTTTACTATGTTACTGTTACCTGCGAGTGTGATTTTAAATAGAGTTTTC 1867
 Db 188 AGGAGGTACTTTACTATGTTACTGTTACCTGCGAGTGTGATTTTAAATAGAGTTTTC 129
 QY 1868 TGCAGTATGCTTTTAGTCTTAAGAAAGCACAGATGGAGCAATCTTCTCTTTTGAAGA 1927
 Db 128 TGCAGTATGCTTTTAGTCTTAAGAAAGCACAGATGGAGCAATCTTCTCTTTTGAAGA 69
 QY 1928 GAATCCCAAAAGTTAGTTTCACTTCTTAAAGTGAATTTGTTTAACTTAAACTGGGCAAC 1987
 Db 58 GAATCCCAAAAGTTAGTTTCACTTCTTAAAGTGAATTTGTTTAACTTAAACTGGGCAAC 9
 QY 1988 TTTTGAAG 1995
 Db 8 TTTTGAAG 1

RESULT 7
 BM473578
 LOCUS BM473578 1182 bp mRNA linear EST 05-PEB-2002
 DEFINITION AGENCOURT 6484439 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5538045
 5', mRNA sequence.
 ACCESSION BM473578
 VERSION BM473578.1 GI:18522620
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

http://image.llnl.gov
Plate: LLNL12230 row: e column: 22
High quality sequence start: 20
High quality sequence stop: 756.

```

Location/Qualifiers
1. .1182
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:5538045"
  /clone.lib="NIH MGC 72"
  /tissue_type="melanotic melanoma"
  /lab_hosts="DHIOB (phage-resistant)"
  /note="Organ: skin; vector: pCMV-Sport6; Site_1: NotI;
  Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
  Average insert size 2 kb. Library constructed by Life

```

Query Match	42.6%;	Score 914.6;	DB 13;	Length 1182;
Best Local Similarity	94.7%;	Pred. No. 2.7e-239;		
Matches 1033; Conservative	0;	Mismatches 46;	Indels 12;	Gaps 8

Qy 92 GGGGCTCAAGGTGTGCATGTGTGACGGGAAGAGAGAGAGAGAGCGCCCTCAGAGGTG 151
Dd 96 GGGGCTCAAGGTGTGCATGTGTGACGGGAAGAGAGAGAGAGAGCGCCCTCAGAGGTG 155			

Qy	212	TAAGGAAGGAGAATCTTAGCAAGCAATGTCTCAAGATGGTGCCTTCAGTTCCAAG	271
Db	216	TAAGGAAGGAGAATCTTAGCAAGCAATGTCTCAAGATGGTGCCTTCAGTTCCAAG	275

Qy 332 CAGCATCATCATGAATTTGAGCACACCAAAAAAGACCTGGATGGATTTTCGAAGCTAT 391
|||||
Db 336 CAGCATCATCATCATGAATTTGAGCACACCAAAAAAGACCTGGATGGATTTTCGAAGCTAT 395
|||||

Qy 452 CTGAAGATTTCGATTCAACCCCTATGAAAAGATAAAGGCCAGGGGCTTCCTCGTGAATAATAT 511
|||||
Db 456 CTGAAGATTTCGATTCAACCCCTATGAAAAGATAAAGGCCAGGGGCTTCCTCGTGAATAATAT 515

Qy	572	GCTGCAAAAGCCCTTAAAGTC	TGATGGTG	TGAGGAATG	AAGAATAC	CTTTCT	TGATCTGA	631
Db	576	GCTGCAAAAGCCCTTAAAGTC	TGATGGTG	TGAGGAATG	AAGAATAC	CTTTCT	TGATCTGA	635

692	TGAACTCTTTTAAACCGGATGAAGATACCAAAAAAATACTACAGAAGTACAACTATTGTC	QV
696	TGAACTCTTTTAAACCGGATGAAGATACCAAAAAAATACTACAGAAGTACAACTATTGTC	Db
752	GTGTGAAATCTCAGCTTTCAATTCAGACAGGTACCCGAGGATTTAATAAAGAATCTTTTAC	QV

Db	816	872	875
QY	TTCTGTAGCAAGGACGTCTTACTCAGGGGAAAATACAGAAAGCTTGATGCCCTCCAG	GTCAATGTCGATTTACGCCAGTTTCTACAACTCTGGATTCGTTGATACC-TTTATAGGA	930

Db	936	GGAGGCCAAGAGNATATTNTTGGTGTCTACATATGATATATCTGGGNTGCCACAGTGGGA	935
Qy	986	-ATCTGPTATTCTTTAATCATCTAATGAACCCACCAATGGAACGC--TGTGAATTTG	1042

1100	ATGAAGGCCAA	1110
1108	TCATGGAGGTCACACCAATATAAAATCCCTCTGCATATGTAAGGGGGGGGAATCCTCCCTCCCA	1111
Qy		
Db		
1116	ATGAAGGCCAA	1126

DEFINITION	AL541422	LTI_FLO02_PL1 Homo sapiens cDNA clone CS0DE006VF01 5 prime
LOCUS	AL541422	926 bp mRNA linear EST 16-FEB-2000
ACCESSION	AL541422	
VERSION	AL541422.1	GI:12872478

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 926)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

```

FEATURES
  source
    1. .926
      Location/Qualifiers
        /organism="Homo sapiens"
        /db_xref="taxon:9606"

```

cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life

Technologies, Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 285 a 176 c 231 g 230 t 4 others

ORIGIN

Query Match 40.9%; Score 878.4; DB 9; Length 926;
Best Local Similarity 99.2%; Pred. No. 28-229;
Matches 909; Conservative 4; Mismatches 0; Indels 3; Gaps 3;

QY 40 GGTGACGGGTGACAAAGGGGGTTAGCAGCTGGGCTGGACCGTTAGCGAGGGGCTCA 99
Db 1 GGTGACGGGTGACAAAGGGGGTTAGCAGCTGGGCTGGACCGTTAGCGAGGGGCTCA 60

QY 100 AGGTGTGCATGTGTGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 159
Db 61 AGGTGTGCATGTGTGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

QY 160 CTGCGAGCTTCTTCCCGGGGGCCATAAAGCCGCCCAATTTCCAGCTGTAAAGGA 219
Db 121 CTGCGAGCTTCTTCCCGGGGGCCATAAAGCCGCCCAATTTCCAGCTGTAAAGGA 180

QY 220 GAGGAAGATCTTAGCAAGCAATGTCTCAAGATGGTCTTCTCAGTTTCCAGAGTCAT 279
Db 181 GAGGAAGATCTTAGCAAGCAATGTCTCAAGATGGTCTTCTCAGTTTCCAGAGTCAT 240

QY 280 CGCAAGAGCTAGATTTATCTGTGAAGAGAGAACTAGAAAAATATCTACACAGATCA 339
Db 241 CGCAAGAGCTAGATTTATCTGTGAAGAGAGAACTAGAAAAATATCTACACAGATCA 300

QY 340 TCATCATGATTTGAGCAGACACCAAAAGAGCTGGATGGATTTCCGAGCTATTCATGA 399
Db 301 TCATCATGATTTGAGCAGACACCAAAAGAGCTGGATGGATTTCCGAGCTATTCATGA 360

QY 400 TTTTTCAGAAAGGGGCTTCTGTGGATTGGGAAAAATCCAGAGACCCCTGAAGAT 459
Db 361 TTTTTCAGAAAGGGGCTTCTGTGGATTGGGAAAAATCCAGAGACCCCTGAAGAT 420

QY 460 TCGATTCAACCCCTATGAAAGAGATAAAGCCAGGGCTTGCCTGATATATATCTTCCG 519
Db 421 TCGATTCAACCCCTATGAAAGAGATAAAGCCAGGGCTTGCCTGATATATATCTTCCG 480

QY 520 TTGAACAAACTAGTGTGTGAACTCAATGTGTGGTGGGACACAGATGGCTGCAAA 579
Db 481 TTGAACAAACTAGTGTGTGAACTCAATGTGTGGTGGGACACAGATGGCTGCAAA 539

QY 580 GGCCTTAAAGTCTGATGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTCA 639
Db 540 GGCCTTAAAGTCTGATGTGAGGAATGAGAAATACCTTTCTGGATCTGACTGTTCA 599

QY 640 CAAATTTGAACATTTGAATAAACCTAACAATACAGATGTTCTCTGTTTTTAATGA 699
Db 600 CAAATTTGAACATTTGAATAAACCTAACAATACAGATGTTCTCTGTTTTTAATGA 659

QY 700 TTTAACCCGATGAAGATACCAAAATACTACAGAGTACATCAATGTCGTGTGAA 759
Db 660 TTTAACCCGATGAAGATACCAAAATACTACAGAGTACATCAATGTCGTGTGAA 719

QY 760 ATCTACACTTCAATCAAGCAGGTACCCGAGGATTAATAAGAACTTTTACTTCCGT 819
Db 720 ATCTACACTTCAATCAAGCAGGTACCCGAGGATTAATAAGAACTTTTACTTCCGT 779

QY 820 GCAAGGAGCTGTCTTACTCAGGGGAAATACAGAGCTTGTGACCTCAGGTCAAT- 878
Db 780 GCAAGGAGCTGTCTTACTCAGGGGAAATACAGAGCTTGTGACCTCAGGTCAATGG 839

QY 879 TGATATTTAGCCAGTTTCTACAACTCTGG-ATTGCTTGTATACCTTTTATAGGAGGCA 937
Db 840 TGATATTTAGCCAGTTTCTACAACTCTGGATTGCTTGTATACCTTTTATAGGAGGCA 909

QY 938 AAGAGTATATTTTGT 953

Db 900 AAGAGTATATTTTGT 915

RESULT 9

AL561196

LOCUS

DEFINITION

AL561196

ACCESSION

AL561196

VERSION

AL561196.1

KEYWORDS

EST.

SOURCE

ORGANISM

Human

REFERENCE

1 (bases 1 to 965)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 965

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DL001YK06"

/clone_lib="LTI_NFL010_BC2"

/sex="male"

/tissue_type="B cells from Burkitt lymphoma"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with NotI and

cloned into the NotI and EcoRV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 313 a 186 c 218 g 245 t 3 others

ORIGIN

Query Match 40.8%; Score 875; DB 9; Length 965;

Best Local Similarity 97.9%; Pred. No. 1.8e-228;

Matches 947; Conservative 3; Mismatches 8; Indels 9; Gaps 6;

QY 125 GAGAGAGAGAGGGGCGC-CTCAGAGTGACCTTTCAGCTGCGAGCTTCTTCCCGGGCG 183

Db 1 GAGAGAGAGAGGGGCGCCTCAGAGTGACCTTTCAGCTGCGAGCTTCTTCCCGGGCG 60

QY 184 CCAATAACGCCCCCAATTTCCAGCTGCTAAAGAGAGAGAGATCTTAGC-AAAGCAAT 242

Db 61 CCAATAACGCCCCCAATTTCCAGCTGCTAAAGAGAGAGAGATCTTAGC-AAAGCAAT 120

QY 243 GTCTCAAGATGGTG----CTTCTCAGTTCCAGAGATCATTCGCAAGAGCTAGATTTAT 298

Db 121 GTCTCAAGATGGTGATATCTCAGTTCCAGAGATCATTCGCAAGAGCTAGATTTAT 180

QY 299 GTGTGAAGAGAGAACTAGAAAAATACTACACACAGCATCATCATGATTTGAGCACA 358

Db 181 GTGTGAAGAGAGAACTAGAAAAATACTACACACAGCATCATCATGATTTGAGCACA 240

QY 359 CCAAAAAAGACCTGGATGGATTTCCGAAAGCTATTTTCATAGATTTTTCGAAAGAGGGGC 418

Db 241 CCAAAAAAGACCTGGATGGATTTCCGAAAGCTATTTTCATAGATTTTTCGAAAGAGGGGC 300

QY 419 CTTCTGTGATTTGGGAAAAATCCAGAGACCCCTCGAGATTCGATTCACACCTATGAA 478

Db 301 CTTCTGTGATTTGGGAAAAATCCAGAGACCCCTCGAGATTCGATTCACACCTATGAA 360

Db 721 AAAAAAATACTACAGAAGTACCATCATTTGCTGGTGGAATACTACACATTTCAATCAAAAG 780
 Qy 781 AGTACCCGAGGATTAATAAGAAATCTTTACTTCTGTAGCAAGAGCGTGTCTTACTCA 840
 Db 781 CAGTACCCGAGGATTAATAAGAAATCTTTACTTCTGTAGCAAGAGCGTGTCTTACTCA 840
 Qy 841 GGGGAAAAATACAGAAGCTGGTACCTCCAGGTCATGGTGATATTATAGCCCA-GTTTCTA 899
 Db 841 AGGG-AAATACAGAAGCTGGTACCTCCAGGTCATGGTGATATTATAGCCCA-GTTTCTA 899
 Qy 900 CAACTCT-GGATGCTTGATACCTTTATAGGAGAGGCAAGA---GTATATTTTGTGT 955
 Db 900 CAACTCTGGGATCGCTGGATACCTTTATAGGAGAGGCAAGA---GTATATTTTGTGT 955
 Qy 956 CTAACATAGATAATCTGGGTGCCACAGTGG 985
 Db 960 CTAACATAGATAATCTGGGTGGCACCGG 989

RESULT 11
 BQ213239 908 bp mRNA linear EST 02-MAY-2002
 LOCUS AGENCOURT 7560136 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6055751
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ213239
 VERSION BQ213239.1 GI:20394234
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL1316 row: h column: 24
 High quality sequence stop: 643.
 Location/Qualifiers
 1..908
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6055751"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

FEATURES

source
 1..908
 300 a 167 c 193 g 248 t
 BASE COUNT
 ORIGIN
 Query Match 39.6%; Score 850.8; DB 14; Length 908;
 Best Local Similarity 97.8%; Pred. No. 7.5e-222;
 Matches 884; Conservative 0; Mismatches 17; Indels 3; Gaps 2;
 Qy 597 TGGTGTGAGGAATGAGATACCTTTCTGGATCTGACTGTTCAGCAAAATTGAACATTGAA 656
 Db 5 TTGGGTGAGGAATGAGATACCTTTCTGGATCTGACTGTTCAGCAAAATTGAACATTGAA 64
 Qy 657 TAAACCTACATACAGATGTTCTCTGTTTAAATGAACCTCTTTTAAACAGGATGAAGA 716
 Db 65 TAAACCTACATACAGATGTTCTCTGTTTAAATGAACCTCTTTTAAACAGGATGAAGA 124

Qy 717 TACCAAAAAAATACTACAGAAGTACAATCATTTGCTGTGAAAAATCTACACTTTCAATCA 776
 Db 125 TACCAAAAAAATACTACAGAAGTACAATCATTTGCTGTGAAAAATCTACACTTTCAATCA 184
 Qy 777 AAGCAGGTACCCGAGGATTAATAAGAAATCTTTACTTCTGTAGCAAGAGCGTGTCTTA 836
 Db 185 AAGCAGGTACCCGAGGATTAATAAGAAATCTTTACTTCTGTAGCAAGAGCGTGTCTTA 244
 Qy 837 CTCAGGGGAAAAATACAGAAGCTTGGTACCTCCAGGTCATGGTGATATTATACGCCAGTTT 896
 Db 245 CTCAGGGGAAAAATACAGAAGCTTGGTACCTCCAGGTCATGGTGATATTATACGCCAGTTT 304
 Qy 897 CTACAACTCTCGATTGCTTGATACCTTTATAGGAGAGGCAAGAGTATATTTTGTGTGTC 956
 Db 305 CTACAACTCTCGATTGCTTGATACCTTTATAGGAGAGGCAAGAGTATATTTTGTGTGTC 364
 Qy 957 TAACATAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCATCTTAATGAACCC 1016
 Db 365 TAACATAGATAATCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCATCTTAATGAACCC 424
 Qy 1017 ACCCAATGGAAGACGCTGTGAATTTGTGCATGGAAGTCAAAATAAAACACGTCGACAGTGT 1076
 Db 425 ACCCAATGGAAGACGCTGTGAATTTGTGCATGGAAGTCAAAATAAAACACGTCGACAGTGT 484
 Qy 1077 AAAGGGCGGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGT 1136
 Db 485 AAAGGGCGGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGGAAATTTGCTCAAGT 544
 Qy 1137 GCCAAAAGCAGATGTAGACAGTTCAGTCTGTATCAAAGTTCAAATATTTAATACAAA 1196
 Db 545 GCCAAAAGCAGATGTAGACAGTTCAGTCTGTATCAAAGTTCAAATATTTAATACAAA 604
 Qy 1197 CAACCTATGGATTTCTCTTGCAGCAGTTTAAAGACTGCAGGAGCAAAATGCCATTGACAT 1256
 Db 605 CAACCTATGGATTTCTCTTGCAGCAGTTTAAAGACTGCAGGAGCAAAATGCCATTGACAT 664
 Qy 1257 GGAATCAATTTGTAATGCAAAAGACTTTTGGATGGAGCCCTGAATGTCAATTTAGAAAC 1316
 Db 665 GGAATCAATTTGTAATGCAAAAGACTTTTGGATGGAGCCCTGAATGTCAATTTAGAAAC 724
 Qy 1317 TGCAGTAGGGGTGCGCATCAAAGTTTGGAGAAATCTCTAGGTATTATATGCGCAAGG-- 1374
 Db 725 TGCAGTAGGGGTGCGCATCAAAGTTTGGAGAAATCTCTAGGTATTATATGAGCCAAAGGA 784
 Qy 1375 AGCGTCTTCTGCTGTCAAAACCAACATCAGATCTCTTGTGGTGATGTCAAAACCTCTAT 1434
 Db 785 GCGGATTACTGCTCTGTCAAAACCAACATCAGATCTCTTGTGGTGATGTCAAAACCTCTAT 844
 Qy 1435 AGTCTTAATGAGGATCTCTGACCAATGAGTGAAAGC-GGGAAATTTCTACAGTCCCTT 1493
 Db 845 AGTCTTAATGAGGATCTCTGACCAATGAGTGAAAGCGGGGATTCTTCTACAGTGGCCT 904
 Qy 1494 GTTT 1497
 Db 905 TGGT 908
 RESULT 12
 AL568945/c
 LOCUS AL568945 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE006YL06 3 prime
 DEFINITION mRNA sequence.
 ACCESSION AL568945
 VERSION AL568945.1 GI:12923792
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 AUTHORS Full-length cDNA libraries and normalization
 TITLE Unpublished (2001)
 JOURNAL

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES Location/Qualifiers
source 1..857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE006Y106"
/clone_lib="LTI FL002_PL1"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 246 a 180 c 144 g 283 t 4 others
ORIGIN

Query Match 39.2%; Score 841.8; DB 9; Length 857;
Best Local Similarity 99.3%; Pred. No. 2.1e-219;
Matches 851; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1037 AATTGTGTCATGGAAGTCAAAATAAAGACGTCAGATGTAAGGCGGGACACTCACTC 1096
DB 857 AATTGTCATGG-AGTCAAAATAAAGACGTCAGATGTAAGGCGGGACACTCACTC 799
QY 1097 AATATGAAGGCAAACTGAGACTGGTGGAAATGCTCAAGTGCCAAAGACATGTAGACG 1156
DB 798 AATATGAAGGCAAACTGAGACTGGTGGAAATGCTCAAGTGCCAAAGACATGTAGACG 739
QY 1157 AGTTCAAGTCTGTATCAAAAGTTCAAAATATTTAATACAAACCTATGGATTTCTCTTG 1216
DB 738 AGTTCAAGTCTGTATCAAAAGTTCAAAATATTTAATACAAACCTATGGATTTCTCTTG 679
QY 1217 CAGCAGTTAAAGACTCGAGGAGCAAAATGCCATTGACATGGAAATCATTTGTAATGCAA 1276
DB 678 CAGCAGTTAAAGACTCGAGGAGCAAAATGCCATTGACATGGAAATCATTTGTAATGCAA 619
QY 1277 AGACTTTGGATGGAGGCTGTAATGTCAATTAAGTAAAGTGCAGTGGGCTGCATCA 1336
DB 618 AGACTTTGGATGGAGGCTGTAATGTCAATTAAGTAAAGTGCAGTGGGCTGCATCA 559
QY 1337 AAAGTTTGGAGATTTCTCTAGGTATTAATGTGCCAAGGAGCGGTTTTCTGCTGTCAAAA 1396
DB 558 AAAGTTTGGAGATTTCTCTAGGTATTAATGTGCCAAGGAGCGGTTTTCTGCTGTCAAAA 499
QY 1397 CCACATCAGATCTCTGTGTTGATGTCAAACCTCTATAGTCTTAATCAGGATCTCTGA 1456
DB 498 CCACATCAGATCTCTGTGTTGATGTCAAACCTCTATAGTCTTAATCAGGATCTCTGA 439
QY 1457 CAATGAGTGAAGGCGGAATTTCTCAGATGCGCTTGGTTAAATTAGCGAGTCTCTTTA 1516
DB 438 CAATGAGTGAAGGCGGAATTTCTCAGATGCGCTTGGTTAAATTAGCGAGTCTCTTTA 379
QY 1517 CGAAGTTTCAAGATTTATCTAAGAAGATTGAAAGTATACAGATATGTTGAATGGATC 1576
DB 378 CGAAGTTTCAAGATTTATCTAAGAAGATTGAAAGTATACAGATATGTTGAATGGATC 319
QY 1577 ACCTCAGAGTTTCAGGAGATGTGACATTTGGAAAAATGTTTCATTAAGGGAACGGTTA 1636
DB 318 ACCTCAGAGTTTCAGGAGATGTGACATTTGGAAAAATGTTTCATTAAGGGAACGGTTA 259
QY 1637 TCATCATTTGCAAAATCATGGTGACAGAAATTTGATATCCACCTGGAGCAGTATTAGAGACA 1696
DB 258 TCATCATTTGCAAAATCATGGTGACAGAAATTTGATATCCACCTGGAGCAGTATTAGAGACA 199
QY 1697 AGATTGTGTCGGAAACCTTCGTCATCTTTGGACCACTGAAATGAAAAATATCTGTGGACACT 1756

DB 198 AGATTGTGTCGGAAACCTTCGTCATCTTGGACCAMARAAATGAAAAATCTGTGGACACT 139
QY 1757 TAAATAATGGCTAGTTTCTTACAAATGAAATGTTCTCTAGGATTTCTAAAATAGGACGTA 1816
DB 138 TAAATAATGGCTAGTTTCTTACAAATGAAATGTTCTCTAGGATTTCTAAAATAGGACGTA 79
QY 1817 CTTTACTATGTTACTGTACCTCGCAGTGTGTTTAAAAATAGAGTCTTCTGCAGTATG 1876
DB 78 CTTTACTATGTTACTGTACCTCGCAGTGTGTTTAAAAATAGAGTCTTCTGCAGTATG 19
QY 1877 CTTTACTGTCTAAGAAAA 1893
DB 18 CTTTACTGTCTAAGAAAA 2

RESULT 13
BQ223934 872 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT 7504319 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6049993
DEFINITION 5', mRNA sequence.
ACCESSION BQ223934
VERSION BQ223934.1 GI:20405334
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 872)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LIAM13301 row: 1 column: 02
High quality sequence stop: 691.
Location/Qualifiers
source 1..872
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6049993"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 280 a 154 c 194 g 244 t
ORIGIN

Query Match 37.7%; Score 810; DB 14; Length 872;
Best Local Similarity 98.5%; Pred. No. 1.1e-210;
Matches 839; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

QY 849 TACAGAAGCTTGGTACCTCCAGGTCACTGGTGATATTACGCCAGTTTCTCAACTCTGG 908
DB 1 TACAGAAGCTTGGTACCTCCAGGTCACTGGTGATATTACGCCAGTTTCTCAACTCTGG 60
QY 909 ATTCTTGTATACCTTTTATAGGAGAGGCAAGAGATATATTTTGTGTCTAATAGATAA 968
DB 61 ATTCTTGTATACCTTTTATAGGAGAGGCAAGAGATATATTTTGTGTCTAATAGATAA 120
QY 969 TCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCCAATGGAAA 1028
DB 121 TCTGGGTGCCACAGTGGATCTGTATATTTCTTAATCATCTAATGAACCCCAATGGAAA 180

```
1029 ACCTGTGTAATTTGTTCATCGAAGTCACAAATAAAACACGTGCAGATGTAAAGCGCGGAC 1088
181 ACCTGTGTAATTTGTTCATCGAAGTCACAAATAAAACACGTGCAGATGTAAAGCGCGGAC 240
1089 ACTCACTCAATATGAAGCAAACTGAGACTGTGTGAAATTTGCTCAAGTGCCAAAGACACA 1148
241 ACTCACTCAATATGAAGCAAACTGAGACTGTGTGAAATTTGCTCAAGTGCCAAAGACACA 300
1149 TGTAGACAGTTCAAGTCTGTATCAAAAGTTCAAAATTTTATATACAAACCACTATGAT 1208
301 TGTAGACAGTTCAAGTCTGTATCAAAAGTTCAAAATTTTATATACAAACCACTATGAT 360
1209 TTCTCTTGACGAGTTAAAGAGCTGCAGGAGCAAAATGCAATTTGATGCAATCAATGAT 1268
361 TTCTCTTGACGAGTTAAAGAGCTGCAGGAGCAAAATGCAATTTGATGCAATCAATGAT 420
1269 GAATGCAAAAGACTTTGGATGGAGGCTGAATGTCATTTCAATTTAGAACTGCAGTAGGGGC 1328
421 GAATGCAAAAGACTTTGGATGGAGGCTGAATGTCATTTCAATTTAGAACTGCAGTAGGGGC 480
1329 TGCATCAAAAGTTTGTGAAATTTCTAGGTATTAATGTGCAAGAGCGGTTTCTGCGC 1388
481 TGCATCAAAAGTTTGTGAAATTTCTAGGTATTAATGTGCAAGAGCGGTTTCTGCGC 540
1389 TGTCAAAACCCATCAGATCTCTTGTGCTGTGATGTCAAACCTCTATAGTCTTAATGCGAG 1448
541 TGTCAAAACCCATCAGATCTCTTGTGCTGTGATGTCAAACCTCTATAGTCTTAATGCGAG 600
1449 ATCTCTGACAATGAGTGAAAGCGGGAATTTCTCTACAGTGCCTTTGGTTAAATAGGCAG 1508
601 ATCTCTGACAATGAGTGAAAGCGGGAATTTCTCTACAGTGCCTTTGGTTAAATAGGCAG 660
1509 TTCTTTTACGAAGTTCAAGATTTCTAAGAAGATTTGAAGTATACAGATATGCTTGA 1568
661 TTCTTTTACGAAGTTTCAAGATTTCTAAGAAGATTTGAAGTATACAGATATGCTTGA 720
1569 ATTGGATCACTCACAGTTTTCAGGAGATGTGACATTTGGAAATAATGTTTCAATAAGGG 1628
721 ATTGGATCACTCACAGTTTTCAGGAGATGTGACATTTGGAAATAATGTTTCAATAAGGG 780
1629 AACGGTTATCATCTGCAAAATCATGG--TGACAGAAATGATATCCACCT--GGAGCAGT 1685
781 AACGGTTATCATCTGCAAAATCATGGGTGACAGAAATTTGATATCCACCTGGGAGCAGT 840
1686 ATTAGAGACAA 1697
841 ATTTTAGAGAAA 852

RESULT 14
AL565147/c
LOCUS AL565147 LTI_Fl015_Brn1 Homo sapiens cDNA clone CSODN004YM14 3
DEFINITION prime, mRNA Sequence.
ACCESSION AL565147
VERSION AL565147.1 GI:12916232
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 823)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
1. .823
Location/Qualifiers
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
/clone="CSODN004YM14"
/clone_lib="LTI_Fl015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was constructed by Life Technologies.
Contact : Feng Liang Life Technologies, a division of
Invitrogen 9800 Medical Center Drive Rockville, Maryland
20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 238 a 169 c 132 g 261 t 23 others
ORIGIN

Query Match 37.1%; Score 797.4; DB 9; Length 823;
Best Local Similarity 96.0%; Pred. No. 3.1e-207;
Matches 790; Conservative 23; Mismatches 10; Indels 0; Gaps 0;

QY 1179 CAAATATTTAATCAAAACCACTATGATTTCTTCTGACAGCTTAAAGACATGCAGGA 1238
Db 823 CAAATATTTAATCAAAACCACTATGATTTCTTCTGACAGCTTAAAGACATGCAGGA 764
QY 1239 GCATAATGCCATTGACATGGAATCATTTGTAATGCAAGACTTTGGATGGAGGCTGAA 1298
Db 763 SCATAATGCCATTGACATGGAATCATTTGTAATGCAAGACTTTGGATGGAGGCTGAA 704
QY 1299 TGTCAATCAATAGAACTGCAGTAGGGCTGCCATCAAAAGTTTTCAGAAATTTCTTAGG 1358
Db 703 TGTCAATCAATAGAACTGCAGTAGGGCTGCCATCAAAAGTTTTCAGAAATTTCTTAGG 644
QY 1359 TATTAAATGTGCAAGAGCGGCTTTTCTGCTGTGCAAAACCAATCAGATCTCTTCTGCTGT 1418
Db 643 TATTAAATGTGCAAGAGCGGCTTTTCTGCTGTGCAAAACCAATCAGATCTCTTCTGCTGT 584
QY 1419 GATGTCAAACTCTATAGTCTTAATGCAAGATCTCTGCAATGAGTGAAGAGCGGAAT 1478
Db 583 GATVCAAACTCTATAGTCTTAATGCAAGATCTCTGCAATGAGTGAAGAGCGGAATK 524
QY 1479 TCCTACAGTGCCTTTGGTTAAATTTAGGCAGTCTTTTACGAAGTTTCAAGATTTATCTAAG 1538
Db 523 TCCTACAGTGCCTTTGGTTAAATTTAGGCAGTCTTTTACGAAGTTTCAAGATTTATCTAAG 464
QY 1539 AAGATTTGAAAGTATACAGATATGCTTGAATTTGATCACCCTCAGATTTTCAAGAGATGT 1598
Db 463 AAGATTTGAAAGTATACAGATATGCTTGAATTTGATCACCCTCAGATTTTCAAGAGATGT 404
QY 1599 GACATTTGGAAATAATGTTTCAATTAAGGGAGCGGTTATCATTTGCAAAATCATGGTGA 1658
Db 403 GASATTTGGAAATAATGTTTCAATTAAGGGAGCGGTTATCATTTGCAAAATCATGGTGA 344
QY 1659 CAGAATTCATATCCACCTGGAGCAGTATTAGAGAAACAAGATTTGTCTGCGAAACCTTCG 1718
Db 343 CAGAATTCATATCCACCTGGAGCAGTATTAGAGAAACAAGATTTGTCTGCGAAACCTTCG 284
QY 1719 CATCTTGACCACTGAAATGAAATACTGTGGACACTTAAATATAGGCTAGTTTCTTA 1778
Db 283 CATCTTGACCACTGAAATGAAATACTGTGGACACTTAAATATAGGCTAGTTTCTTA 224
QY 1779 CAATGAATGTTCTCTAGGATTTTAAATAGCAGGTTACTTTTACTATGTTACTGTACCT 1838
Db 223 CAATGAATGTTCTCTAGGATTTTAAATAGCAGGTTACTTTTACTATGTTACTGTACCT 164
QY 1839 GCAGTGTGATTTTAAATAGAGTTTCTGCAAGTATGCTTTTACTCTAAGAAAGACACA 1898
Db 163 GCAGTGTGATTTTAAATAGAGTTTCTGCAAGTATGCTTTTACTCTAAGAAAGACACA 104
QY 1899 GATGAGCAATCTTTCTCTTTGAAGAGATCCCAAGTTAGTTTCTATCTTAAGTGC 1958
Db 103 GATGAGCAATCTTTCTCTTTGAAGAGATCCCAAGTTAGTTTCTATCTTAAGTGC 44
```

```

QY 1959 AATATTGTTTAAATCTTAAACTGGCAACTTTGGAAGAACTTT 2001
|||||
Db 43 AATATTGTTTAAATCTTAAACTGGCCACKTGTGAAGACGTTT 1

RESULT 15
AUI19264
LOCUS HEMBA1005415 845 bp mRNA linear EST 01-AUG-2002
DEFINITION HEMBA1005415 5', mRNA
ACCESSION AUI19264
VERSION AUI19264
KEYWORDS GI:10934499
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 845)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
    source
        1..845
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="HEMBA1005415"
            /clone_lib="HEMBA1"
            /tissue_type="whole embryo, mainly head"
            /dev_stage="embryo, 10 weeks"
            /note="Vector: pME18SFL3"
BASE COUNT 259 a 161 c 221 g 200 t 4 others
ORIGIN
    Query Match 36.9%; Score 791.6; DB 9; Length 845;
    Best Local Similarity 98.2%; Pred. No. 1.2e-205;
    Matches 830; Conservative 0; Mismatches 12; Indels 3; Gaps 3;
QY 3 GTTTAGGAGAAAGTAGGGGCTGTGGGTGTCGGGAGCGCGCTGACGGGTGGACAAGGGGG 62
|||||
Db 1 GTTTAGGAGAAAGTAGGGGCTGTGGGTGTCGGGAGCGCGCTGACGGGTGGACAAGGGGG 60
|||||
QY 63 GTTAGCAGCTGGGCTGCGACCGTTAGGGAGGGGCTCAAGGTGTGCATGTGTGAGGGAAGA 122
|||||
Db 61 GTTAGCAGCTGGGCTGCGACCGTTAGGGAGGGGCTCAAGGTGTGCATGTGTGAGGGAAGA 120
|||||
QY 123 GAGAGAGAGAGAGGGCGCTCAGAGGTGACCTTCAGCCCTGCGAGCGCTTCTCCCGGGGC 182
|||||
Db 121 GAGAGAGAGAGAGGGCGCTCAGAGGTGACCTTCAGCCCTGCGAGCGCTTCTCCCGGGGC 180
|||||
QY 183 GCCATAAAGCCGCCCAATTTCCAGCTGCTAAAGGAAGAGAGATCTTAGCAAAAGCAAT 242
|||||
Db 181 GCCATAAAGCCGCCCAATTTCCAGCTGCTAAAGGAAGAGAGATCTTAGCAAAAGCAAT 240
|||||
QY 243 GTCTCAAGATGTGTCTTCTCAGTCCAGAAGTCATTCCGCAAGAGCTAGAAATTATCTGT 302
|||||
Db 241 GTCTCAAGATGTGTCTTCTCAGTCCAGAAGTCATTCCGCAAGAGCTAGAAATTATCTGT 300
|||||
QY 303 GAAGAAGGAAGTAAGAAAAAATACTACCAAGATCATCATGAATTTGAGCACACCAA 362
|||||

```

Search completed: June 28, 2003, 08:15:44
Job time : 1991.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:03:25 ; Search time 5626.03 Seconds
(without alignments)
17018.784 Million cell updates/sec

Title: US-09-981-353-172

Perfect score: 3290

Sequence: 1 cggcctctctcgtggttc.....agtttttagaaaaaaa 3290

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sta.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3283.6	99.8	3297	9	BC019944	BC019944 Homo sapi
2	3270.8	99.4	3328	6	AX3281702	AX3281702 Sequence
3	3175.2	96.5	3220	6	AX330185	AX330185 Sequence
4	3175.2	96.5	3220	6	AX330695	AX330695 Sequence
5	3175.2	96.5	3220	6	AX330914	AX330914 Sequence
6	3175.2	96.5	3220	6	HUNGVLVR1X	L20859 Human leuke
7	3172	96.4	3211	6	II11794	II11794 Sequence 1
8	2166	65.8	3228	10	BC015085	BC015085 Mus muscu
9	2133	64.8	3227	6	AX305391	AX305391 Sequence
10	2133	64.8	3227	10	MUSGLVR1	M73696 Murine Glvr
11	2133	64.8	3260	6	II11795	II11795 Sequence 3
12	1954	59.4	2890	10	AB000489	AB000489 Rattus no
13	1934.4	58.8	2040	10	CGU13946	U13946 Cricetus
14	1838.6	55.9	2128	4	AF074085	AF074085 Felis cat
15	1820.2	55.3	2503	10	AF063024	AF063024 Cricetus
16	960.8	29.2	140678	9	AC079922	AC079922 Homo sapi
17	960.8	29.2	178531	2	AC067802	AC067802 Homo sapi
18	946.2	28.8	203269	2	AC080187	AC080187 Homo sapi
19	904.4	27.5	200000	9	AB019437	AB019437 Homo sapi
20	901.6	27.4	241393	2	AC127381	AC127381 Homo sapi
21	772	23.5	810	11	G06485	G06485 human STS-W
22	707.2	21.5	3178	9	HSLGLVR1G4	AF102062 Homo sapi
23	624.4	19.0	2429	10	AF063025	AF063025 Cricetus
24	617.2	18.8	3175	6	AX329634	AX329634 Sequence
25	617.2	18.8	3175	6	I25452	I25452 Sequence 1
26	617.2	18.8	3175	6	I43897	I43897 Sequence 1
27	617.2	18.8	3175	9	HUNGVLVR2X	L20852 Human leuke
28	617.2	18.8	3690	9	BC028600	BC028600 Homo sapi
29	616.8	18.7	1959	10	AF239675	AF239675 Cricetus
30	615.8	18.7	1962	4	AF394194	AF394194 Felis cat
31	615.2	18.7	1959	10	CGU13945	U13945 Cricetus
32	615	18.7	2287	10	RATRAM1A	L19931 Rat amphotr
33	602	18.3	2747	9	HSLGLVR1G1	AF102059 Homo sapi
34	601.8	18.3	2251	10	AF196476	AF196476 Mus muscu
35	598	18.2	200000	9	AB019440	AB019440 Homo sapi
36	550.8	16.7	1194	10	AF172628S4	AF172631 Mus muscu
37	550.8	16.7	251670	2	AL772347	AL772347 Mus muscu
38	511	15.5	150425	2	AC121664	AC121664 Rattus no
39	507.8	15.4	196516	2	AL844218	AL844218 Mus muscu
40	481	14.6	175022	2	AC103019	AC103019 Rattus no
41	448.8	13.6	1148	10	AF172628S7	AF172634 Mus muscu
42	448	13.6	495	6	AX397089	AX397089 Sequence
43	440.4	13.4	175022	2	AC103019	AC103019 Rattus no
44	421.8	12.8	2842	10	AF172628S1	AF172628 Mus muscu
45	420.8	12.8	158902	2	AC013778	AC013778 Homo sapi

ALIGNMENTS

RESULT 1
BC019944
LOCUS BC019944
DEFINITION Homo sapiens, clone MGC:8767 IMAGE:3918690, mRNA, linear PRI 22-JAN-2002
ACCESSION BC019944
VERSION BC019944.1 GI:18044776
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3297)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCRP/DBP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: ang@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 26 Row: b Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7382462.

FEATURES

source

1..3297
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:8767 IMAGB:3918690"
 /issue_type="Skin, melanotic melanoma."
 /clone_lib="NIH MGC_72"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 447..2486
 /codon_start=1
 /product="Unknown (protein for MGC:8767)"
 /protein_id="AAH19944.1"
 /db_xref="GI:18044777"
 /translation="MATLITSTTAATAASGLVDVLMILGLIFIAFLVAFSGVANDV
 ANSGTAVSGSVTLKQACILASIFETVGSVLLGAKVSETIRKGLIDVEMYNSTOGLL
 MAGSVAMGSAVQVAFSLKPLISGTHICIVGATIGFSLVANGQGVKWSLILKIVM
 SNFVPLLSGIMGILFVLVAFILHKADVPNGLRALFVACTVGINLFIMYTG
 PLIGFDKPLMWGLTILSVCAVFCALIVFVCPMKRIEREIKCSPSEPLMEKN
 SLKDEHETKLSVGDLENKHPVSEVGPATVQAQVEERTVSFKLGDLEAEPERLPL
 SVDLKEETISDTPNCAVOLPNGNLVQFSAVSNQINSSGHVQYHTVHKDSLGLYKELL
 HKLHLAKGDCWDGSKPLRNNNSYVTMAICMPLDSFRAKGEOKGSEMEKLTW
 PNADSKRIRMSQSYCNASDLHSASEIDMSVKAEMGLGRKSGNSLEWYDQDK
 PEVSLLFQLQILITACGFAHGGNDVNAIGLVALYLYDIDVSSKVAIPIMLLL
 YGGVGI CVGLWVNGRVIQTMGKDLPTIPSSGFSIELASALTIVIASNIGLPISTTH
 CKVGSVSVGLRSKKAQVDRFLFRNIFMAFVTPISGVISAIMAIFRYILRM"

CDS

797 a 732 c 785 g 983 t
 BASE COUNT
 ORIGIN

Query Match 99.8%; Score 3283.6; DB 9; Length 3297;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3286; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CGCGCTCTCTGCGTGGTCTTCTTCTCGCCCGCTGAACCCCGCGGTCTTCTGGG 60
 DB 4 CGCGCTCTCTGCGTGGTCTTCTTCTCGCCCGCTGAACCCCGCGGTCTTCTGGG 63
 QY 61 AAGGTGCTGAGTCCCGCTGAGTGTCTCCCGTCCGCGAGCCCGCGGTCTTCTGGG 120
 DB 64 AAGGTGCTGAGTCCCGCTGAGTGTCTCCCGTCCGCGAGCCCGCGGTCTTCTGGG 123
 QY 121 CTCACGCGCTCGCGCTCGATCTCTCGCTCTCCCGCTCGCGCTCTTCTTCTTCTGGAT 180
 DB 124 CTCACGCGCTCGCGCTCGATCTCTCGCTCTCCCGCTCGCGCTCTTCTTCTTCTGGAT 183
 QY 181 GAATCTGCTCTTCTCTCTCGCGCTCGATCTCTCGCTCTTCTTCTTCTTCTTCTTCTG 240
 DB 184 GAATCTGCTCTTCTCTCTCTCGCGCTCGATCTCTCGCTCTTCTTCTTCTTCTTCTTCTG 243
 QY 241 AGCCAAAGAAACCCAGACAAACAGATGCCCATACGACGCGGTATAGCAGTAACCTCCCCAGC 300

244 AGCCAAAGAAACCCAGACAAACAGATGCCCATACGACGCGGTATAGCAGTAACCTCCCCAGC 303
 QY 301 TCGGTTTCTGTCGCGTAGTCTTACAGTATTTTATATATATATATATATATATATATATATAT 360
 DB 304 TCGGTTTCTGTCGCGTAGTCTTACAGTATTTTATATATATATATATATATATATATATATAT 363
 QY 361 GCATTTTGTATACCTCATATCTGTTTACACATCTTGAAGGCGCTCAGTAGTCTCTTCTTA 420
 DB 364 GCATTTTGTATACCTCATATCTGTTTACACATCTTGAAGGCGCTCAGTAGTCTCTTCTTA 423
 QY 421 CTAAACAAACCACTACTCCAGAGAAATGGCAACGGTGATTAACCACTACTACAGTCTGTACCG 480
 DB 424 CTAAACAAACCACTACTCCAGAGAAATGGCAACGGTGATTAACCACTACTACAGTCTGTACCG 483
 QY 481 CGGCTTCTGTCGCTTCTGTTGGTACCTACTATGATGCTCATCTCGGCTTCTTATTTGAT 540
 DB 484 CGGCTTCTGTCGCTTCTGTTGGTACCTACTATGATGCTCATCTCGGCTTCTTATTTGAT 543
 QY 541 TTGCTCTTGGCATTTCTCCGTGGGAGCAATGATGTAGCAAAATCTTTTGGTACAGCTGTGG 600
 DB 544 TTGCTCTTGGCATTTCTCCGTGGGAGCAATGATGTAGCAAAATCTTTTGGTACAGCTGTGG 603
 QY 601 GCTCAGGTGTAGTGACCCCTGAGCAAGCCTGCATCCTAGCTAGCATCTTTTGAACAGTGG 660
 DB 604 GCTCAGGTGTAGTGACCCCTGAGCAAGCCTGCATCCTAGCTAGCATCTTTTGAACAGTGG 663
 QY 661 GCTCTGCTTACTGGGGGCAAGTGAGGAAACCATCGGAGGCTTGTATTGAGCTGG 720
 DB 664 GCTCTGCTTACTGGGGGCAAGTGAGGAAACCATCGGAGGCTTGTATTGAGCTGG 723
 QY 721 AGATGTACAACTCGACTCAAGGGCTACTGATGGCGGCTCAGTCACTGCTATGTTGGTT 780
 DB 724 AGATGTACAACTCGACTCAAGGGCTACTGATGGCGGCTCAGTCACTGCTATGTTGGTT 783
 QY 781 CTGCTGTGTGGCAACTCGTGGCTTCGTTTTTGAAGCTCCCTATTTCTGGAACCCATTTGA 840
 DB 784 CTGCTGTGTGGCAACTCGTGGCTTCGTTTTTGAAGCTCCCTATTTCTGGAACCCATTTGA 843
 QY 841 TTGTTGGTGCACACTATTGGTTCTCCCTCGTGGCAAGGGGAGGAGGTGTCAAGTGT 900
 DB 844 TTGTTGGTGCACACTATTGGTTCTCCCTCGTGGCAAGGGGAGGAGGTGTCAAGTGT 903
 QY 901 CTGAACATGATAAAAAATGTGTCTTGGTCTCGTCCCACTGCTTTCTGGAATATATGT 960
 DB 904 CTGAACATGATAAAAAATGTGTCTTGGTCTCGTCCCACTGCTTTCTGGAATATATGT 963
 QY 961 CTGGAATTTTATTTCTTCTGTTCTGTCATTCCTCCATTAAGGAGATCCAGTTCCTTA 1020
 DB 964 CTGGAATTTTATTTCTTCTGTTCTGTCATTCCTCCATTAAGGAGATCCAGTTCCTTA 1023
 QY 1021 ATGGTTTGGGAGCTTGGCAGTTTCTATGCTGACAGTGGGAATAAACCTCTTTTCCA 1080
 DB 1024 ATGGTTTGGGAGCTTGGCAGTTTCTATGCTGACAGTGGGAATAAACCTCTTTTCCA 1083
 QY 1081 TCATGTATATCTGAGCACCGTTGCTGGGCTTTTGACAACTTCTCTGTTGGGTACCATCC 1140
 DB 1084 TCATGTATATCTGAGCACCGTTGCTGGGCTTTTGACAACTTCTCTGTTGGGTACCATCC 1143
 QY 1141 TCATCTCGTGGGATGTGCAGTTTCTGTGCCCTTATCGTCTGGTTCTTTGTATGTCCCA 1200
 DB 1144 TCATCTCGTGGGATGTGCAGTTTCTGTGCCCTTATCGTCTGGTTCTTTGTATGTCCCA 1203
 QY 1201 GGATGAAGAAAAATTTGAACGAGAAATAAGTGTAGTCTTCTGAAAGCCCTTAATGG 1260
 DB 1204 GGATGAAGAAAAATTTGAACGAGAAATAAGTGTAGTCTTCTGAAAGCCCTTAATGG 1263
 QY 1261 AAAAAAGAAATAGCTTGAAGAAAGACCATGAAGAAACAAAGTTGCTGTTGGTGATTTG 1320
 DB 1264 AAAAAAGAAATAGCTTGAAGAAAGACCATGAAGAAACAAAGTTGCTGTTGGTGATTTG 1323
 QY 1321 AAAAAAGCAATCCTGTTCTGAGGTAGGCGCTGCCACATGTGCCCTTCCAGGCTGTGGTGG 1380

Db 1324 AAAACAGCATCTGTTTCTGAGGTAGGGCTGCCACTGTGCCCTCCAGGCTGTGGTG 1383
Qy 1381 AGGAGAGAACAGTCTCATTTCAAACTTTGGAGATTGGAGGAAGCTCCAGAGAGAGAGGC 1440
Db 1384 AGGAGAGAACAGTCTCATTTCAAACTTTGGAGATTGGAGGAAGCTCCAGAGAGAGAGGC 1443
Qy 1441 TTCCAGCGTGACCTTGAAGAGAGAAACCAGCATAGATAGCACCCTGTAATGTGTGCGTGC 1500
Db 1444 TTCCAGCGTGACCTTGAAGAGAGAAACCAGCATAGATAGCACCCTGTAATGTGTGCGTGC 1503
Qy 1501 AGTTGCTTAATGGGAACCTTTGCTCAGTTCCAGTCAAGCCGTCAAGCAACCAAAATAAACTCCA 1560
Db 1504 AGTTGCTTAATGGGAACCTTTGCTCAGTTCCAGTCAAGCCGTCAAGCAACCAAAATAAACTCCA 1563
Qy 1561 GTGGCCACTACCAAGTATACACCGTGCATAAAGATTCCGGCCCTGTACAAAGAGACTACTCC 1620
Db 1564 GTGGCCACTACCAAGTATACACCGTGCATAAAGATTCCGGCCCTGTACAAAGAGACTACTCC 1623
Qy 1621 ATAAATTACATCTTGCCAAAGTGGGAGATTGCATGGGAGACTCCGGTGACAAACCTTTAA 1680
Db 1624 ATAAATTACATCTTGCCAAAGTGGGAGATTGCATGGGAGACTCCGGTGACAAACCTTTAA 1683
Qy 1681 GCGCAATAATAGCTATACCTCTATACCATTGSCAATATGTGGCATGCGCTCTGGATTCTAT 1740
Db 1684 GCGCAATAATAGCTATACCTCTATACCATTGSCAATATGTGGCATGCGCTCTGGATTCTAT 1743
Qy 1741 TCCGTGCCAAAGAGTGAACAGAGGGCGAAGAAATGGAGAGAGCTGACATGCCCTTAATG 1800
Db 1744 TCCGTGCCAAAGAGTGAACAGAGGGCGAAGAAATGGAGAGAGCTGACATGCCCTTAATG 1803
Qy 1801 CAGACTCCAAGAGCGAATTCGAATGGACAGTTTACACAGTTACTGCAATGCTGTGTCG 1860
Db 1804 CAGACTCCAAGAGCGAATTCGAATGGACAGTTTACACAGTTACTGCAATGCTGTGTCG 1863
Qy 1861 ACCTTCACTCAGCATCTGAGATAGACATCAGTGTCAAGCAGAGATGGTCTAGGTGACA 1920
Db 1864 ACCTTCACTCAGCATCTGAGATAGACATCAGTGTCAAGCAGAGATGGTCTAGGTGACA 1923
Qy 1921 GAAAGGAAGTAAATGGCTCTCTAGAGAAATGGTATGATGACCAAGATAGCCCTGAAGTCTCTC 1980
Db 1924 GAAAGGAAGTAAATGGCTCTCTAGAGAAATGGTATGATGACCAAGATAGCCCTGAAGTCTCTC 1983
Qy 1981 TCCTCTTCAGTTCTCTGAGATGCTTTACAGCTGCTTTGGGTCAATTCGCCCATGTTGGGCA 2040
Db 1984 TCCTCTTCAGTTCTCTGAGATGCTTTACAGCTGCTTTGGGTCAATTCGCCCATGTTGGGCA 2043
Qy 2041 ATGAGCTAAGCAATGCCATTTGGGCTCTGTTGCTTTATATTTGGTTATGACACAGGAG 2100
Db 2044 ATGAGCTAAGCAATGCCATTTGGGCTCTGTTGCTTTATATTTGGTTATGACACAGGAG 2103
Qy 2101 ATGTTTCTTTCAAAAGTGGCAACCAATATGGCTTCTACTATGTTGGTGTGTTGTTATCT 2160
Db 2104 ATGTTTCTTTCAAAAGTGGCAACCAATATGGCTTCTACTATGTTGGTGTGTTGTTATCT 2163
Qy 2161 GTTGTGCTCTGTTGGTGGGAGAGATTATCCAGACCATGTTGGGAGAGATCTGACAC 2220
Db 2164 GTTGTGCTCTGTTGGTGGGAGAGATTATCCAGACCATGTTGGGAGAGATCTGACAC 2223
Qy 2221 CGATCACACCTCTAGTGGCTTCAGTATTGAACTGGCATCTGCCCTCACTGTGGTGTG 2280
Db 2224 CGATCACACCTCTAGTGGCTTCAGTATTGAACTGGCATCTGCCCTCACTGTGGTGTG 2283
Qy 2281 CATCAAAATATGGCCCTTCCCATCAGTACCAACATTTGAAAGTGGGCTCTGTTGTGTCG 2340
Db 2284 CATCAAAATATGGCCCTTCCCATCAGTACCAACATTTGAAAGTGGGCTCTGTTGTGTCG 2343
Qy 2341 TTGGCTGGCTCCGGTCCAAGAGGCTGTTGACTGGCGTCTCTTTGCTAAATTTTATGG 2400
Db 2344 TTGGCTGGCTCCGGTCCAAGAGGCTGTTGACTGGCGTCTCTTTGCTAAATTTTATGG 2403
Qy 2401 CTTGGTTGCTCAGTCCCAATTTCTGGAGTTATCAGTCTGCCATCATGGCAATCTTCA 2460
Db 2404 CTTGGTTGCTCAGTCCCAATTTCTGGAGTTATCAGTCTGCCATCATGGCAATCTTCA 2463

Qy 2461 GATATGTCTATCCTCAGAAATGTGAAGCTGTGTTGAGATTAAATTTGTGTCAATTTGTTGGGA 2520
Db 2464 GATATGTCTATCCTCAGAAATGTGAAGCTGTGTTGAGATTAAATTTGTGTCAATTTGTTGGGA 2523
Qy 2521 CCATCTTAGGTATTCCTGCTCCCTGCTGAAGAATGATTACAGTGTAAACAGAAAGACTGACAA 2580
Db 2524 CCATCTTAGGTATTCCTGCTCCCTGCTGAAGAATGATTACAGTGTAAACAGAAAGACTGACAA 2583
Qy 2581 GAGTCTTTTATTTGGGAGCCAGAGAGGGAGTGTACTTGTGCTATAACTGCTTTTCT 2640
Db 2584 GAGTCTTTTATTTGGGAGCCAGAGAGGGAGTGTACTTGTGCTATAACTGCTTTTCT 2643
Qy 2641 GCTAAATATGAATTCCTCAAAATTAGCTGTGTGTAAATAGCCCGGTTCCACTGGCTCT 2700
Db 2644 GCTAAATATGAATTCCTCAAAATTAGCTGTGTGTAAATAGCCCGGTTCCACTGGCTCT 2703
Qy 2701 GCTGAGGTCCTCTTCTCTGGGCTGTGAATTCCTGTACATATTTCTACTATTTTGA 2760
Db 2704 GCTGAGGTCCTCTTCTCTGGGCTGTGAATTCCTGTACATATTTCTACTATTTTGA 2763
Qy 2761 TCAGGCTTCAATTCATTTATTTAAATGTTGCTCTGAAAGATGACTTGTGATTTTTTT 2820
Db 2764 TCAGGCTTCAATTCATTTATTTAAATGTTGCTCTGAAAGATGACTTGTGATTTTTTT 2823
Qy 2821 TCTTTTTTAAACCATGAAGAGCCGTTTGACAGAGCATGCTCTGCTGTTGTTGGTTTCA 2880
Db 2824 TCTTTTTTAAACCATGAAGAGCCGTTTGACAGAGCATGCTCTGCTGTTGTTGGTTTCA 2883
Qy 2881 CAGCTTCGCTCTACATGCACAGGGATTTAAACAAACAAATATATACATCTCCCTT 2940
Db 2884 CAGCTTCGCTCTACATGCACAGGGATTTAAACAAACAAATATATACATCTCCCTT 2943
Qy 2941 GTAGTCTCTTATATAAGTAGAGTCTCTGCTACTCTGCCCTCTGTCAGTAGTGGCAGGAT 3000
Db 2944 GTAGTCTCTTATATAAGTAGAGTCTCTGCTACTCTGCCCTCTGTCAGTAGTGGCAGGAT 3003
Qy 3001 CTATTGGCATATTCGGGAGCTTCTTTAGAGGATGAGTTCTTTGAACACAGTGAAAAATTT 3060
Db 3004 CTATTGGCATATTCGGGAGCTTCTTTAGAGGATGAGTTCTTTGAACACAGTGAAAAATTT 3063
Qy 3061 AAATAGTAACTTTTTCGAAGCAGTTTATGACATGTTATGCTGAAGAAGATGAAGAA 3120
Db 3064 AAATAGTAACTTTTTCGAAGCAGTTTATGACATGTTATGCTGAAGAAGATGAAGAA 3123
Qy 3121 GAAAGAGCTGTGTGCAATCTTTGTTATTTCTTTAAAGATTTCTGCGAGTGTGGGATGAT 3180
Db 3124 GAAAGAGCTGTGTGCAATCTTTGTTATTTCTTTAAAGATTTCTGCGAGTGTGGGATGAT 3183
Qy 3181 GAATGAAGTGAATGTGAACCTTTGGGCAAGTTTAAATGGGACAGCTTCCATGTTTCAATTTG 3240
Db 3184 GAATGAAGTGAATGTGAACCTTTGGGCAAGTTTAAATGGGACAGCTTCCATGTTTCAATTTG 3243
Qy 3241 TCTACTCTTAACTGAATTAAGAGCCCTACAGTTTATAGAAAAA 3290
Db 3244 TCTACTCTTAACTGAATTAAGAGCCCTACAGTTTATAGAAAAA 3293

RESULT 2

AX281702 LOCUS AX281702 3328 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 111 from Patent WO0177389.
ACCESSION AX281702
VERSION AX281702.1 GI:16608953
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Shiffman, D., Somogyi, R., Lawn, R., Seilhamer, J.J., Porter, G.J.,
Mikita, T. and Tai, J.
AUTHORS Genes expressed in foam cell differentiation
TITLE

JOURNAL Patent: WO 0177389-A 111 18-OCT-2001;

Incyte Genomics, Inc. (US)
Location/Qualifiers
1. .3328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 474592.3"

BASE COUNT 792 a 744 c 802 g 990 t

ORIGIN

Query Match 99.4%; Score 3270.8; DB 6; Length 3328;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3283; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY	1	CGCGCTCTCTGGTGGTCTTCTTCTCGCGCGCTGAACCCCGCGGCTGCTTCTGGG	60
DB	34	CGCGCTCTCTGGTGGTCTTCTTCTCGCGCGCTGAACCCCGCGGCTGCTTCTGGG	93
QY	61	AAGTCTGAGTCCCGCTCAGCTGTCCCGGTGCCCGACCCGGCCGCTGTGCCGTGG	120
DB	94	AAGTCTGAGTCCCGCTCAGCTGTCCCGGTGCCCGACCCGGCCGCTGTGCCGTGG	153
QY	121	CTCCAGCCGCTCGCGCTCGATCTCTCTCTCCCGCTCCCGCTCCCGCTTCCCTGGAT	180
DB	154	CTCCAGCCGCTCGCGCTCGATCTCTCTCTCCCGCTCCCGCTTCCCTGGAT	213
QY	181	GAACTTGGCTCTTCTCTCTCCCGCATGGAATCTGCTCGTGTCTTTAGCCCTCTG	240
DB	214	GAACTTGGCTCTTCTCTCTCCCGCATGGAATCTGCTCGTGTCTTTAGCCCTCTG	273
QY	241	AGCCAAAGAAACCCAGACAAACAGATGCCATACGAGCGGTATAGCAGTAATCTCCCGACG	300
DB	274	AGCCAAAGAAACCCAGACAAACAGATGCCATACGAGCGGTATAGCAGTAATCTCCCGACG	333
QY	301	TCGGTTCTGTGCGGTGTTACAGTAATTAATTTTATATAATATAATATAATATAATATA	360
DB	334	TCGGTTCTGTGCGGTGTTACAGTAATTAATTTTATATAATATAATATAATATAATATA	393
QY	361	GCATTTTTCATACCTCATATCTGTTTACATCTTGAAGGGCTCAGTAGTCTCTTA	420
DB	394	GCATTTTTCATACCTCATATCTGTTTACATCTTGAAGGGCTCAGTAGTCTCTTA	453
QY	421	CTAAACAAACACTACTCCAGAGAAATGGCAACGGTGATTAACAGTACTACAGCTGTACCG	480
DB	454	CTAAACAAACACTACTCCAGAGAAATGGCAACGGTGATTAACAGTACTACAGCTGTACCG	513
QY	481	CGCTTCTGCTCTTGTGGGACTACTATGAGTGTCTATCTCGGCTTCAATTTGCAAT	540
DB	514	CGCTTCTGCTCTTGTGGGACTACTATGAGTGTCTATCTCGGCTTCAATTTGCAAT	573
QY	541	TTGCTTTGGCATCTCCGTGGGAGCAATGATGTAGCAAAATCTTTTGTACAGCTGTGG	600
DB	574	TTGCTTTGGCATCTCCGTGGGAGCAATGATGTAGCAAAATCTTTTGTACAGCTGTGG	633
QY	601	GCTCAGGTGTAGTGAACCTTGAACCAAGCTGCACTCTAGTAGCATCTTTTGAACAGTGG	660
DB	634	GCTCAGGTGTAGTGAACCTTGAACCAAGCTGCACTCTAGTAGCATCTTTTGAACAGTGG	693
QY	661	GCTCTGTCTTACTTGGGGGCAAAAGTGAAGCAACATCCGGAAGGGCTTGAATGACGTGG	720
DB	694	GCTCTGTCTTACTTGGGGGCAAAAGTGAAGCAACATCCGGAAGGGCTTGAATGACGTGG	753
QY	721	AGATGTACAACTCGACTCAAGGGCTACTGATGGCCGGCTCAGTCAGTGTCTTGTGGTT	780
DB	754	AGATGTACAACTCGACTCAAGGGCTACTGATGGCCGGCTCAGTCAGTGTCTTGTGGTT	813
QY	781	CTGCTGTGTGGCAACTCGTGGCTTCTGTTTGAAGCTTCCCTATTTCTGGAACCCATTGTA	840
DB	814	CTGCTGTGTGGCAACTCGTGGCTTCTGTTTGAAGCTTCCCTATTTCTGGNAACCCATTGTA	873
QY	841	TTGTTGGTCAACTATTTGTTTCTCCCTCGTGCAAAAGGGGCAAGGGGTGTCAGTGGT	900
DB	874	TTGTTGGTCAACTATTTGTTTCTCCCTCGTGCAAAAGGGGCAAGGGGTGTCAGTGGT	933

QY	901	CTGAACATGATAAAATTTGATGTCCTGTTTCGTTCCTCCCACTGCTTCTTCTGAAATATGT	960
DB	934	CTGAACATGATAAAATTTGATGTCCTGTTTCGTTCCTCCCACTGCTTCTTCTGAAATATGT	993
QY	961	CTGGAATTTTATCTTCTCTGTTTCGTGCAATTCATCTCCATTAAGGACAGATCCAGTTCCTA	1020
DB	994	CTGGAATTTTATCTTCTCTGTTTCGTGCAATTCATCTCCATTAAGGACAGATCCAGTTCCTA	1053
QY	1021	ATGTTTTCGGAGCTTTCGACGTTTCTATGCTGACAGATTTGGAATAACCTCTTTTCCA	1080
DB	1054	ATGTTTTCGGAGCTTTCGACGTTTCTATGCTGACAGATTTGGAATAACCTCTTTTCCA	1113
QY	1081	TCATGTATACCTGGAGCACCGTTGCTGGGCTTTTGACAAATCTTCTCTGTTGGGTACCATCC	1140
DB	1114	TCATGTATACCTGGAGCACCGTTGCTGGGCTTTTGACAAATCTTCTCTGTTGGGTACCATCC	1173
QY	1141	TCATCTCGGTGGGATGTGACGTTTCTGTGCCCCTTATCGTCTGGTCTTTTGTATGCCA	1200
DB	1174	TCATCTCGGTGGGATGTGACGTTTCTGTGCCCCTTATCGTCTGGTCTTTTGTATGCCA	1233
QY	1201	GGATGAAGAGAAAATTTGAACGAGAAATAAAGTGTAGTCTTCTGAAAGCCCTTAAATGG	1260
DB	1234	GGATGAAGAGAAAATTTGAACGAGAAATAAAGTGTAGTCTTCTGAAAGCCCTTAAATGG	1293
QY	1261	AAAAAAGAAATAGCTTTGAAGAGAACCATGAAGAAACAAAGTTGCTGTTGGTGATATTG	1320
DB	1294	AAAAAAGAAATAGCTTTGAAGAGAACCATGAAGAAACAAAGTTGCTGTTGGTGATATTG	1353
QY	1321	AAAAAAGCAATCTGTTTCTGAGGTAGGGCTGCCACATGTGCCCCCTCCAGGCTGTGGTG	1380
DB	1354	AAAAAAGCAATCTGTTTCTGAGGTAGGGCTGCCACATGTGCCCCCTCCAGGCTGTGGTG	1413
QY	1381	AGGAGAGAAACAGTCTCATTTCAAACTTGAGATTTTGGAGAGAGCTCCAGAGAGAGAGGC	1440
DB	1414	AGGAGAGAAACAGTCTCATTTCAAACTTGAGATTTTGGAGAGAGCTCCAGAGAGAGAGGC	1473
QY	1441	TTCCAGAGCTGGACCTTGAAGAGGAAACAGCATAGATAGCACCGTGAATGGTGCAGTGC	1500
DB	1474	TTCCAGAGCTGGACCTTGAAGAGGAAACAGCATAGATAGCACCGTGAATGGTGCAGTGC	1533
QY	1501	AGTTCCCTTAATGGGAACTTGTCCAGTTTCAGTCAAGCCGTGAGCAACCAAAATAAACTCA	1560
DB	1534	AGTTCCCTTAATGGGAACTTGTCCAGTTTCAGTCAAGCCGTGAGCAACCAAAATAAACTCA	1593
QY	1561	GTGGCCACTACAGTATCACACCGTGCATAGGATTCGGGCTGTACAAAGAGCTACTCC	1620
DB	1594	GTGGCCACTACAGTATCACACCGTGCATAGGATTCGGGCTGTACAAAGAGCTACTCC	1653
QY	1621	ATAAATTTACATCTTCCCAAGGTGGGAGATTGATGGGAGACTCCCGGTGACAAACCTTAA	1680
DB	1654	ATAAATTTACATCTTCCCAAGGTGGGAGATTGATGGGAGACTCCCGGTGACAAACCTTAA	1713
QY	1681	GGCGCAATTAATAGCTATATCTTCTATACCATGGCAATATGTGGCATGCTCTGGATTCA	1740
DB	1714	GGCGCAATTAATAGCTATATCTTCTATACCATGGCAATATGTGGCATGCTCTGGATTCA	1773
QY	1741	TCCGTGCCAAGAAGGTGAACAGAGGGGGAAGAAATGGAGAGCTGATGAGCCCTAATG	1800
DB	1774	TCCGTGCCAAGAAGGTGAACAGAGGGGGAAGAAATGGAGAGCTGATGAGCCCTAATG	1833
QY	1801	CAGACTCCAAAGAGGAAATTCGAATGGACAGTTTACACCAAGTTTACTGCAATGCTGTCTG	1860
DB	1834	CAGACTCCAAAGAGGAAATTCGAATGGACAGTTTACACCAAGTTTACTGCAATGCTGTCTG	1893
QY	1861	ACCTTCTACTCAGCATCTCAGATAGACATGAGTGTCAAGGACAGATGGGTCTAGGTGACA	1920
DB	1894	ACCTTCTACTCAGCATCTCAGATAGACATGAGTGTCAAGGACAGATGGGTCTAGGTGACA	1953
QY	1921	GAAAGGAGTAATGGCTCTCTAGAGAAATGGTATGACAGGATTAAGCTGAAGTCTCTC	1980
DB	1954	GAAAGGAGTAATGGCTCTCTAGAGAAATGGTATGACAGGATTAAGCTGAAGTCTCTC	2013

1981 TCCTCTCCAGTTCCTGCGAGATCCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGTTGGCA 2040
Db TCCTCTTCAGTTCCTGCGAGATCCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGTTGGCA 2073
Qy ATG-ACGTAAGCAATGCCATTCGGGCTCTGGTTCCTTTATATTTGGTTTATGACACAGGA 2099
Db ATGACGTAAGCAATGCCATTCGGGCTCTGGTTCCTTTATATTTGGTTTATGACACAGGA 2133
Qy GATGTTTCTTCAAAAGTGGCAACACCAATATGCTTCTACTCTATGTTGTTGTTATC 2159
Db GATGTTTCTTCAAAAGTGGCAACACCAATATGCTTCTACTCTATGTTGTTGTTATC 2193
Qy TGTGTTGCTGCTGGGTTGGGAGAGAGATTCACAGCATGCGGAGAGATCTGACA 2219
Db TGTGTTGCTGCTGGGTTGGGAGAGAGATTCACAGCATGCGGAGAGATCTGACA 2253
Qy CCGATCACACCTCTAGTGGCTTCAGTATTGAATGAGCTGCGCTCACTGCTGTTGATT 2279
Db CCGATCACACCTCTAGTGGCTTCAGTATTGAATGAGCTGCGCTCACTGCTGTTGATT 2313
Qy GCATCAAAATATTGGCTTCCCATCAGTACACACATTTGTAAGTGGGCTCTGTTGTTCT 2339
Db GCATCAAAATATTGGCTTCCCATCAGTACACACATTTGTAAGTGGGCTCTGTTGTTCT 2373
Qy GTTGCTGGCTCCGCTCCAGAGGCTGTTGACTGCGCTCTTTTCGTAACATTTTATG 2399
Db GTTGCTGGCTCCGCTCCAGAGGCTGTTGACTGCGCTCTTTTCGTAACATTTTATG 2433
Qy GCTGCTGTTGTCACAGTCCCATTTCTGGAGTTATCAGTGCCTGCGCATCATGCAATCTTC 2459
Db GCTGCTGTTGTCACAGTCCCATTTCTGGAGTTATCAGTGCCTGCGCATCATGCAATCTTC 2493
Qy AGATATGTCATCCTCAGATGTAAGCTGTTTGAGATTAATAATTTGTTGTAATTTTGGG 2519
Db AGATATGTCATCCTCAGATGTAAGCTGTTTGAGATTAATAATTTGTTGTAATTTTGGG 2553
Qy ACCATCTTAGTATTCTGCTCCCTGAGAGATTAACAGTGTGTTAAACAGAGATCTGACA 2579
Db ACCATCTTAGTATTCTGCTCCCTGAGAGATTAACAGTGTGTTAAACAGAGATCTGACA 2613
Qy AGAGTCTTTTATTTGGGAGCAGAGGAGGAGTGTACTTGTGCTATACTGCTTTTGG 2639
Db AGAGTCTTTTATTTGGGAGCAGAGGAGGAGTGTACTTGTGCTATACTGCTTTTGG 2673
Qy TGCTAAATATGAATGTCTCAAAATTTAGCTGTGTAAATAGCCGGGTTCCACTGGCTCC 2699
Db TGCTAAATATGAATGTCTCAAAATTTAGCTGTGTAAATAGCCGGGTTCCACTGGCTCC 2733
Qy TGCTGAGTCCCTTCTGCTGCGCTGTGAATTCCTGTACATATTTCTACTTTTGT 2759
Db TGCTGAGTCCCTTCTGCTGCGCTGTGAATTCCTGTACATATTTCTACTTTTGT 2793
Qy ATCAGCTTCAATTCATATGTTTAAATGTTGTTGAGAGATGCTGCGTGTGTTTCA 2819
Db ATCAGCTTCAATTCATATGTTTAAATGTTGTTGAGAGATGCTGCGTGTGTTTCA 2853
Qy TTTCTTTTAAACCATGAAGCGCTTTGACAGAGATGCTGCGTGTGTTTCA 2879
Db TTTCTTTTAAACCATGAAGCGCTTTGACAGAGATGCTGCGTGTGTTTCA 2913
Qy CCAGCTTCTGCTCAGATGCAAGGATTTAAACAAAAATATAAATACAACTTCCT 2939
Db CCAGCTTCTGCTCAGATGCAAGGATTTAAACAAAAATATAAATACAACTTCCT 2973
Qy TGTAGTCTCTTATATAGTAGTCTTGGTACTCTGCGCTCTGCTAGTAGTGCAGGA 2999
Db TGTAGTCTCTTATATAGTAGTCTTGGTACTCTGCGCTCTGCTAGTAGTGCAGGA 3033
Qy TCTATTGGCATATTGGGAGCTTCTTAGAGGATGAGGTTCTTTGAAACAGTGAATAAT 3059
Db TCTATTGGCATATTGGGAGCTTCTTAGAGGATGAGGTTCTTTGAAACAGTGAATAAT 3093
Qy TAAATTAGTAACTTTTTCGAAGCAGTTTTATTGACTGTTTATTGCTTAAGAAAGTAAGAA 3119

3094 TAAATTAGTAACCTTTTTCGAAGCAGTTTTATTGACTGTTATTGCTAAGAAGTAAGAA 3153
Qy AGAAAAAGCCCTGTTGGCAATCTTGGTATTCTTTTAAAGATTCTGCGCAGTGTGGATGA 3179
Db AGAAAAAGCCCTGTTGGCAATCTTGGTATTCTTTTAAAGATTCTGCGCAGTGTGGATGA 3213
Qy TGAATGAAGTGGAAATGTGAACCTTTGGGCAAGTTAAATGGGACAGCTTCCCATGTTT 3239
Db TGAATGAAGTGGAAATGTGAACCTTTGGGCAAGTTAAATGGGACAGCTTCCCATGTTT 3273
Qy GTCTACCTTCTTAACTGAATATAAAGAGCTACAGTTTGTAGAAAAA 3285
Db GTCTACCTTCTTAACTGAATATAAAGAGCTACAGTTTGTAGAAAAA 3319

RESULT 3
AX330185
LOCUS AX330185 3220 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 694 from Patent WO0194629.
ACCESSION AX330185
VERSION AX330185.1 GI:18103163
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Young P.E., Augustus M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, J.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 694 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
1..3220
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 783 a 710 c 763 g 964 t
ORIGIN

Query Match 96.5%; Score 3175.2; DB 6; Length 3220;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3182; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 97 CCGACCCGCGCGCTGTCGCCGTCCAGCGCTCCAGCGCTCGATCTCCTCGTCTCCG 156
Db 24 CCGGCGCTGCGCTGTCGCCGTCCAGCGCTCCAGCGCTCGATCTCCTCGTCTCCG 83
Qy 157 CTCGCGCTCCTCTTTTCCCTGGATGAACCTTGGCTCTTTTCTCTTCTCGCCCATGGAATTC 216
Db 84 CTCGCGCTCCTCTTTTCCCTGGATGAACCTTGGCTCTTTTCTCTTCTCGCCCATGGAATTC 143
Qy 217 TGCTCCGCTGCTTTTAGCCCTCTGAGCCAAAGAAACCCAGACAAACAGATGCCATAGCC 276
Db 144 TGCTCCGCTGCTTTTAGCCCTCTGAGCCAAAGAAACCCAGACAAACAGATGCCATAGCC 203
Qy 277 AGCGTATAGCAGTAACCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336
Db 204 AGCGTATAGCAGTAACCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 263
Qy 337 ATATAATATATATATTTTATTATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTT 396
Db 264 ATATAATATATATATTTTATTATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTT 323
Qy 397 GAAAGGCCCTCAGTAGTCTCTTACTAAACCAACTACTCCAGAGATGGCAACGGTGA 456
Db 324 GAAAGGCCCTCAGTAGTCTCTTACTAAACCAACTACTCCAGAGATGGCAACGGTGA 383
Qy 457 TTACAGTACTACAGTCTGCTACCGCGCTTCTGGTCTTTGGTGGACTACCTATGATGC 516
Db 384 TTACAGTACTACAGTCTGCTACCGCGCTTCTGGTCTTTGGTGGACTACCTATGATGC 443

QY 517 TCATCTCGGCTTCATTATTGTCATTTGCTTGGCAATTCCTCGTGGGAGCCAAATGATAG 576
DB 444 TCATCTCGGCTTCATTATTGTCATTTGCTTGGCAATTCCTCGTGGGAGCCAAATGATAG 503
QY 577 CAAATTTCTTTTGGTACAGCTGTGGGCTCAGAGTGTAGTGACCTCGAAGCAAGCCTGCATCC 636
DB 504 CAAATTTCTTTTGGTACAGCTGTGGGCTCAGAGTGTAGTGACCTCGAAGCAAGCCTGCATCC 563
QY 637 TAGCTAGCATCTTTGAAACAGTGGGCTGTCTTACTGGGGGCCAAAGTGAGCGAAACCA 696
DB 564 TAGCTAGCATCTTTGAAACAGTGGGCTGTCTTACTGGGGGCCAAAGTGAGCGAAACCA 623
QY 697 TCCGGAGGGCTTGATTGACGTGGAGATGTAACAATCGACTCAAGGGCTACTGATGGCCG 756
DB 624 TCCGGAGGGCTTGATTGACGTGGAGATGTAACAATCGACTCAAGGGCTACTGATGGCCG 683
QY 757 GCTCAGTCAAGTGTATGTTGGTTCCTGCTGTGGCAACTCGTGGCTTCGTTTTGAAGC 816
DB 684 GCTCAGTCAAGTGTATGTTGGTTCCTGCTGTGGCAACTCGTGGCTTCGTTTTGAAGC 743
QY 817 TCCCTATTTCTGGAACCCATTTGTTGGTGCAACTATTTGGTTTCTCCCTCGTGGCA 876
DB 744 TCCCTATTTCTGGAACCCATTTGTTGGTGCAACTATTTGGTTTCTCCCTCGTGGCA 803
QY 877 AGGGCAGGAGGTGTCAGTGTGTAAGTCTGATAAATAATTTGATGTTCTGTTCTGTGT 936
DB 804 AGGGCAGGAGGTGTCAGTGTGTAAGTCTGATAAATAATTTGATGTTCTGTTCTGTGT 863
QY 937 CCCCAGTCTTTCTGGAATTTATGTCGGAATTTTATTTCTCTGGTTCGTGCAATTCATCC 996
DB 864 CCCCAGTCTTTCTGGAATTTATGTCGGAATTTTATTTCTCTGGTTCGTGCAATTCATCC 923
QY 997 TCATTAAGCAGATCCAGTTCCTAATGGTTTGGAGCTTTGGCAGTTTCTATGCTGCA 1056
DB 924 TCATTAAGCAGATCCAGTTCCTAATGGTTTGGAGCTTTGGCAGTTTCTATGCTGCA 983
QY 1057 CAGTTGGAATAACCTCTTTTCCATCATGTAATCTGAGCAGCTTTGGGCTTTGACA 1116
DB 984 CAGTTGGAATAACCTCTTTTCCATCATGTAATCTGAGCAGCTTTGGGCTTTGACA 1043
QY 1117 AACTCTCTGTGGGTACCATCTCTCATCTCGGTGGGATGTGCAAGTTTCTGTGCCCTTA 1176
DB 1044 AACTCTCTGTGGGTACCATCTCTCATCTCGGTGGGATGTGCAAGTTTCTGTGCCCTTA 1103
QY 1177 TGTCTGGTTCTTTGATGTCCAGGATGAAGAGAAAATTTGAAACGAGAAATAAGTGTA 1236
DB 1104 TGTCTGGTTCTTTGATGTCCAGGATGAAGAGAAAATTTGAAACGAGAAATAAGTGTA 1163
QY 1237 GTCCCTCTGAAAGCCCTTAAATGGAAGAAAAGAAATAGCTTGAAGAGAACCATGAAGAA 1296
DB 1164 GTCCCTCTGAAAGCCCTTAAATGGAAGAAAAGAAATAGCTTGAAGAGAACCATGAAGAA 1223
QY 1297 CAAAGTTGTCTGTGGTGATATTGAAACCAAGCATCTCTGTTCTGAGGTAGGGCCCTGCCA 1356
DB 1224 CAAAGTTGTCTGTGGTGATATTGAAACCAAGCATCTCTGTTCTGAGGTAGGGCCCTGCCA 1283
QY 1357 CTGTGCCCTCTCAGGCTGTGGTGGAGAGAGAAACAGTCTCATCTCAAACTTGGAGATTTGG 1416
DB 1284 CTGTGCCCTCTCAGGCTGTGGTGGAGAGAGAAACAGTCTCATCTCAAACTTGGAGATTTGG 1343
QY 1417 AGGAAGCTCCAGAGAGAGAGGCTTTCCAGCTGGACTTGAAGAGAGAACCCAGCATAG 1476
DB 1344 AGGAAGCTCCAGAGAGAGAGGCTTTCCAGCTGGACTTGAAGAGAGAACCCAGCATAG 1403
QY 1477 ATAGCACCGTGAAATGGTCAGTGCAGTTGCCTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1536
DB 1404 ATAGCACCGTGAAATGGTCAGTGCAGTTGCCTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1463
QY 1537 CCGTCAAGCAACCAATAAATCCAGTGGCCACTACCAAGTATCAACCCGTGCATAAGGATT 1596
DB 1464 CCGTCAAGCAACCAATAAATCCAGTGGCCACTACCAAGTATCAACCCGTGCATAAGGATT 1523
QY 1597 CCGGCTGTGTACAAAGAGCTACTCCATAAATTACATCTTGCCAAAGGTGGGAGATTGCATGG 1656

DB 1524 CCGGCTGTGTACAAAGAGCTACTCCATAAATTACATCTTGCCAAAGGTGGGAGATTGCATGG 1583
QY 1657 GAGACTCCGGTGAACAAACCCCTTAAGGGGCAATAATAGCTATATCTTCTATACCATGGCAA 1716
DB 1584 GAGACTCCGGTGAACAAACCCCTTAAGGGGCAATAATAGCTATATCTTCTATACCATGGCAA 1643
QY 1717 TATGTGGCATCGCTCTGGATTTCATTTCCGTGCCAAAGAGAGGTGAACAGAGGGCGGAAGAAA 1776
DB 1644 TATGTGGCATCGCTCTGGATTTCATTTCCGTGCCAAAGAGAGGTGAACAGAGGGCGGAAGAAA 1703
QY 1777 TGGAGAACTGCACATGGCTTAATGCAGACTCCAAGAACGAATTCGAAATGGACAGTTTACA 1836
DB 1704 TGGAGAACTGCACATGGCTTAATGCAGACTCCAAGAACGAATTCGAAATGGACAGTTTACA 1763
QY 1837 CCAGTTACTCAATGCTGTGTGCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGCA 1896
DB 1764 CCAGTTACTCAATGCTGTGTGCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGCA 1823
QY 1897 AGGCAGAGATGGGTCTAGGTGACAGAAAGAGAGTAATGGCTCTCTAGAAGATGATG 1956
DB 1824 AGGCAGAGATGGGTCTAGGTGACAGAAAGAGAGTAATGGCTCTCTAGAAGATGATG 1883
QY 1957 ACCGAGTAAGCTGAAGTCTCTCTCTCTTCCAGTTTCTCAGATCCTTACAGCCTGCT 2016
DB 1884 ACCGAGTAAGCTGAAGTCTCTCTCTCTTCCAGTTTCTCAGATCCTTACAGCCTGCT 1943
QY 2017 TTGGGTCAATTCGCCCATGGTGGCAATGACGTAAGCAATGCCATTTGGGCTCTGTGTGCTT 2076
DB 1944 TTGGGTCAATTCGCCCATGGTGGCAATGACGTAAGCAATGCCATTTGGGCTCTGTGTGCTT 2003
QY 2077 TATATTTGGTTTATGACACAGAGATGTTTCTTCAAAAGTGGCAACCAATATGGCTTC 2136
DB 2004 TATATTTGGTTTATGACACAGAGATGTTTCTTCAAAAGTGGCAACCAATATGGCTTC 2063
QY 2137 TACTCTATGTGGTGTGGTATCTGTGTGGTCTGTGGGTTTGGGGAAGAGAGTTATCC 2196
DB 2064 TACTCTATGTGGTGTGGTATCTGTGTGGTCTGTGGGTTTGGGGAAGAGAGTTATCC 2123
QY 2197 AGACATCGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTCACTATTTGAACCTGG 2256
DB 2124 AGACATCGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTCACTATTTGAACCTGG 2183
QY 2257 CATCTGCCCTCACCTGTGGTATGTCATCAAAATATTGGGCTTTCCCATCAGTACACACATT 2316
DB 2184 CATCTGCCCTCACCTGTGGTATGTCATCAAAATATTGGGCTTTCCCATCAGTACACACATT 2243
QY 2317 GTAAAGTGGGCTCTGTGTGTCTGTGGCTCCGCTCCAAGAGGCTGTGTGACTGGC 2376
DB 2244 GTAAAGTGGGCTCTGTGTGTCTGTGGCTCCGCTCCAAGAGGCTGTGTGACTGGC 2303
QY 2377 GTCTCTTTCTGTAACATTTTATTTAGGCTGGTGTGACAGTCCCATTTCTGGAGTTATCA 2436
DB 2304 GTCTCTTTCTGTAACATTTTATTTAGGCTGGTGTGACAGTCCCATTTCTGGAGTTATCA 2363
QY 2437 GTGCTGCCATCATGSCAATCTTCAATATGTCATCTCAGATGTGAAGCTGTTTGGAGAT 2496
DB 2364 GTGCTGCCATCATGSCAATCTTCAATATGTCATCTCAGATGTGAAGCTGTTTGGAGAT 2423
QY 2497 TAAAAATTTGTCTCAATGTTTGGGACCATCTTGTAGGTATTTCTGCTCCCTCGAAGATGATT 2556
DB 2424 TAAAAATTTGTCTCAATGTTTGGGACCATCTTGTAGGTATTTCTGCTCCCTCGAAGATGATT 2483
QY 2557 ACAGTGTAAACAGAGACTGACAGAGCTTTTATTTTGGGAGCCAGAGGAGGAGTGT 2616
DB 2484 ACAGTGTAAACAGAGACTGACAGAGCTTTTATTTTGGGAG - CAGAGGAGGAGGAGTGT 2542
QY 2617 TACTTGTCTATAACCTGCTTTTGTCTAAATATGAATTTGTCTCAAAATTTAGCTGTGTA 2676
DB 2543 TACTTGTCTATAACCTGCTTTTGTCTAAATATGAATTTGTCTCAAAATTTAGCTGTGTA 2602
QY 2677 ATAGCCCGGTTTCCACTGGCTCTGCTGAGGTCCCTTTTCTTCTTGGGCTGTGAATTCCT 2736

1104	TCGTCCTGGTTCCTTTGTATGTGTCCTCCAGGATGAAGAGAAAAATTGAAACGAGAAAAATTGAACGAGAAATTAAGTGTGA	1163
1237	GTCCTCTCTGAAGAGCCCTTAAATGAAAAAAGAAATAGCTTTGAAAGAAAGACCATCAAGAAAA	1296
1164	GTCCTTCTGAAGAGCCCTTAAATGAAAAAAGAAATAGCTTTGAAAGAAAGACCATCAAGAAAA	1223
1297	CAAAAGTTGTCTGTTGGTGATATTGAAAAACAAGCAATCTGTTTCTGAGGTAGGCGCTGCGCA	1356
1224	CAAAAGTTGTCTGTTGGTGATATTGAAAAACAAGCAATCTGTTTCTGAGGTAGGCGCTGCGCA	1283
1357	CTGTGCCCCCTCCAGGCTGTGGTGAGGAGAGACAGTCTCATTTCAAACCTTTGGAGATTTGG	1416
1284	CTGTGCCCCCTCCAGGCTGTGGTGAGGAGAGACAGTCTCATTTCAAACCTTTGGAGATTTGG	1343
1417	AGGAAGCTCCAGAGAGAGAGAGGCTTCCACGCTGGGACTTTGAAAGAGAAAAACAGCATAG	1476
1344	AGGAAGCTCCAGAGAGAGAGAGGCTTCCACGCTGGGACTTTGAAAGAGAAAAACAGCATAG	1403
1477	ATAGCACCGTGAATTGGTGCAGTGCAGTTGCTTAATGGGAAACCTTTGTCCAGTTTCAGTCAAG	1536
1404	ATAGCACCGTGAATTGGTGCAGTGCAGTTGCTTAATGGGAAACCTTTGTCCAGTTTCAGTCAAG	1463
1537	CCGTGAGCAACCAATAAATCTCAGTGGCCACTACAGATATCACACCGTGCATTAAGGATT	1596
1464	CCGTGAGCAACCAATAAATCTCAGTGGCCACTCCAGATATCACACCGTGCATTAAGGATT	1523
1597	CCGCGCTGTACAAAGAGCTACTCCATAAATTTACATCTTTGCCAAGGTGGGAGATTTGCATGG	1656
1524	CCGCGCTGTACAAAGAGCTACTCCATAAATTTAGATCTTTGCCAAGGTGGGAGATTTGCATGG	1583
1657	GAGACTCCGGTGACAAACCCCTTAAGGGGCAATAATAGCTATPACTTCTTACCATGGCAA	1716
1584	GAGACTCCGGTGACAAACCCCTTAAGGGGCAATAATAGCTATPACTTCTTACCATGGCAA	1643
1717	TATGTGGCATGCCCTCTGATTTCAATTCGGTGCACAAAGAGGTGAACAAGAGGGCGAAGAAA	1776
1644	TATGTGGCATGCCCTCTGATTTCAATTCGGTGCACAAAGAGGTGAACAAGAGGGCGAAGAAA	1703
1777	TGGAGAGCTGACATGGCCTTAATCGACAGTCCCAAGAGCGCAATTCGAATTCGACAGATTACA	1836
1704	TGGAGAGCTGACATGGCCTTAATCGACAGTCCCAAGAGCGCAATTCGAATTCGACAGATTACA	1763
1837	CCAGTTACTGCAATGCTGTGTCTGACACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA	1896
1764	CCAGTTACTGCAATGCTGTGTCTGACACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA	1823
1897	AGGCAGAGATGGGTCTAGTGACAGAAAAAGGAATATGGCTCTCTAGAGAAATGGTATG	1956
1824	AGGCAGAGATGGGTCTAGTGACAGAAAAAGGAATATGGCTCTCTAGAGAAATGGTATG	1883
1957	ACCAGGATAGCCTGAGTCTCTCTCTCTTCAGTTTCTGCGAGATCCTTACAGCCTGCT	2016
1884	ACCAGGATAGCCTGAGTCTCTCTCTCTTCAGTTTCTGCGAGATCCTTACAGCCTGCT	1943
2017	TTGGTCAATTGCCCATGGTGGCAATGACGTAAGCAATGCCATTTGGGCTCTGGTTGCTT	2076
1944	TTGGTCAATTGCCCATGGTGGCAATGACGTAAGCAATGCCATTTGGGCTCTGGTTGCTT	2003
2077	TATATTTGGTTTATGACACAGGAGATGTTTCTTCAAAGTGGCAACCAATATGGCTTC	2136
2004	TATATTTGGTTTATGACACAGGAGATGTTTCTTCAAAGTGGCAACCAATATGGCTTC	2063
2137	TACTCTATGGTGGTGTGGTATCTGTGTGTGTCTGTGGGTTTGGGGAAGAGAGTTATCC	2196
2064	TACTCTATGGTGGTGTGGTATCTGTGTGTGTCTGTGGGTTTGGGGAAGAGAGTTATCC	2123
2197	AGACCATGGGAAGGATCTGACACCGATACACACCTCTAGTGGCTTCAGTATTGCACTGG	2256
2124	AGACCATGGGAAGGATCTGACACCGATACACACCTCTAGTGGCTTCAGTATTGCACTGG	2183
2257	CATCTGCCCTCAGTGTGGTGAATGCATCAAAATATTGGCCCTTCCCATCAGTACAAACATTT	2316
2184	CATCTGCCCTCAGTGTGGTGAATGCATCAAAATATTGGCCCTTCCCATCAGTACAAACATTT	2243

[illegible]

DEFINITION	Sequence 1423 from Patent WO0194629.	
ACCESSION	AX330914	
VERSION	AX330914.1	GI:18103893
KEYWORDS	human.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G., Horrigan, S., Soppet, D. R. and Weaver, Z.	
TITLE	Cancer gene determination and therapeutic screening using signature gene sets	
JOURNAL	Patent: WO 0194629-A 1423 13-DEC-2001;	
FEATURES	Avalon Pharmaceuticals (US)	
source	Location/Qualifiers	
	1..3220	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
BASE COUNT	783 a 710 c 763 g 964 t	
ORIGIN		
	Query Match 96.5%; Score 3175.2; DB 6; Length 3220;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 3192; Conservative 0; Mismatches 6; Indels 1; Gaps 1;	
QY	97 CCGACCCGGCGCGTGGCGGCTCCAGCGCGCTCGATCTCTCGTCTCCCG 156	
DB		
QY	157 CTCGGCCCTCCCTTTCCCTGATGAACCTTGCGTCTCTCTCTCCGCCATGGAATTC 216	
DB		
QY	84 CTCGGCCCTCCCTTTCCCTGATGAACCTTGCGTCTCTCTCTCCGCCATGGAATTC 143	
DB		
QY	217 TGTCTCGTCTTTTAGCCCTCTAGCCAAAGAAACCCAGACAAACAGATGCCATACGC 276	
DB		
QY	144 TGCTCGTCTTTTAGCCCTCTAGCCAAAGAAACCCAGACAAACAGATGCCATACGC 203	
DB		
QY	277 AGCGTATAGCAGTAACCTCCCGCTCGGTTCTGTGCGCGTAGTTTACAGTATTAATTTT 336	
DB		
QY	204 AGCGTATAGCAGTAACCTCCCGCTCGGTTCTGTGCGCGTAGTTTACAGTATTAATTTT 263	
DB		
QY	337 ATATAAT 396	
DB		
QY	264 ATATAAT 323	
DB		
QY	397 GAAAGCGCTCAGTAGTCTCTTACTAAACACCACTACTCCAGAGAATGGCAACGGTGA 456	
DB		
QY	324 GAAAGCGCTCAGTAGTCTCTTACTAAACACCACTACTCCAGAGAATGGCAACGGTGA 383	
DB		
QY	457 TTACCACTACTACAGTGTACCGCGCTTCTGGTCTTTGGTGGAATGAGTATGATGC 516	
DB		
QY	384 TTACCACTACTACAGTGTACCGCGCTTCTGGTCTTTGGTGGAATGAGTATGATGC 443	
DB		
QY	517 TCATCTGGGCTTCATATATGATTTGTCTGGCATTCCTCGTGGAGCAATGATGATAG 576	
DB		
QY	444 TCATCTGGGCTTCATATATGATTTGTCTGGCATTCCTCGTGGAGCAATGATGATAG 503	
DB		
QY	577 CAAATTTCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCTGAGCAAGCTGCATCC 636	
DB		
QY	504 CAAATTTCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCTGAGCAAGCTGCATCC 563	
DB		
QY	637 TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCCAAAGTGAGCGAAACCA 696	
DB		
QY	564 TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCCAAAGTGAGCGAAACCA 623	
DB		
QY	697 TCCGGAAGGCTTTGATTTGACGTGGAGATGTAACTCGACTCAAGGGCTACTGATGGCGG 756	
DB		
QY	624 TCCGGAAGGCTTTGATTTGACGTGGAGATGTAACTCGACTCAAGGGCTACTGATGGCGG 683	
DB		
QY	757 GCTCAGTCACTGCTATGTTGGTCTGTGCTGTGGCAACTCGTGGCTTCTGTTTGAAGC 816	
DB		
QY	684 GCTCAGTCACTGCTATGTTGGTCTGTGCTGTGGCAACTCGTGGCTTCTGTTTGAAGC 743	
DB		

QY	817 TCCCTATTCTGGAAACCCATTGTATTGTTGGTGCAACTATTGGTTTCTCCCTCGTGCAA 876	
DB		
QY	744 TCCCTATTCTGGAAACCCATTGTATTGTTGGTGCAACTATTGGTTTCTCCCTCGTGCAA 803	
DB		
QY	877 AGGGCAGGAGGGTCTCAAGTGGTCTGAACATGATATAAAATTTGTGATGCTTGGTCTGTG 936	
DB		
QY	804 AGGGCAGGAGGGTCTCAAGTGGTCTGAACATGATATAAAATTTGTGATGCTTGGTCTGTG 863	
DB		
QY	937 CCCCACTGCTTCTGGAATTTATGCTGGAAATTTATTTCTTCTCTGTTCTGTCGATTCATCC 996	
DB		
QY	864 CCCCACTGCTTCTGGAATTTATGCTGGAAATTTATTTCTTCTGTTCTGTCGATTCATCC 923	
DB		
QY	997 TCCATAAGGCAGATCCAGTTCCTAATGTTGCGAGCTTTGCGAGTTCCTATGCTGCA 1056	
DB		
QY	924 TCCATAAGGCAGATCCAGTTCCTAATGTTGCGAGCTTTGCGAGTTCCTATGCTGCA 983	
DB		
QY	1057 CAGTTGGAAATAAACCTCTTTTCCATCATGTATACCTGGAGCACCGTTCGTGGGCTTTGACA 1116	
DB		
QY	984 CAGTTGGAAATAAACCTCTTTTCCATCATGTATACCTGGAGCACCGTTCGTGGGCTTTGACA 1043	
DB		
QY	1117 AACTTCTCTGTGGGTACCATCTCATCTCGGTGGATGTGCAGTTTTCTGTGCCCTTA 1176	
DB		
QY	1044 AACTTCTCTGTGGGTACCATCTCATCTCGGTGGATGTGCAGTTTTCTGTGCCCTTA 1103	
DB		
QY	1177 TCGTCTGGTCTTTGTATGTCCAGGATGAAGAGAAAAATTAAGACGAGAAAAATAAGTGTA 1236	
DB		
QY	1104 TCGTCTGGTCTTTGTATGTCCAGGATGAAGAGAAAAATTAAGACGAGAAAAATAAGTGTA 1163	
DB		
QY	1237 GTCTTCTGAAAGCCCTTAAATGGAAGAAAAAGAAATAGCTTGAAGAGAACCATGAAGAAA 1296	
DB		
QY	1164 GTCTTCTGAAAGCCCTTAAATGGAAGAAAAAGAAATAGCTTGAAGAGAACCATGAAGAAA 1223	
DB		
QY	1297 CAAAGTGTCTGTGTGATATTGAAACAAAGCATCTGTCTTCTGAGGTAGGCGCTGCCA 1356	
DB		
QY	1224 CAAAGTGTCTGTGTGATATTGAAACAAAGCATCTGTCTTCTGAGGTAGGCGCTGCCA 1283	
DB		
QY	1357 CTGTGCCCTCCAGGCTGTGTGTGAGAGAGAAACAGTCTCATTTCAAACTTTGGAGATTTGG 1416	
DB		
QY	1284 CTGTGCCCTCCAGGCTGTGTGTGAGAGAGAAACAGTCTCATTTCAAACTTTGGAGATTTGG 1343	
DB		
QY	1417 AGGAAGCTCCAGAGAGAGAGAGGCTTCCAGGCTGGACTTGAAGAGAGAAACAGCATAG 1476	
DB		
QY	1344 AGGAAGCTCCAGAGAGAGAGAGGCTTCCAGGCTGGACTTGAAGAGAGAAACAGCATAG 1403	
DB		
QY	1477 ATAGCACCGTGAATGTCGAGTGCAGTTGCCTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1536	
DB		
QY	1404 ATAGCACCGTGAATGTCGAGTGCAGTTGCCTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1463	
DB		
QY	1537 CCGTCAGCAACCAATAAACTCCAGTGGCCACTACAGTATCACACCGTGCAATAGGATTT 1596	
DB		
QY	1464 CCGTCAGCAACCAATAAACTCCAGTGGCCACTCCAGTATCACACCGTGCAATAGGATTT 1523	
DB		
QY	1597 CCGGCTGTACAAAGAGCTACTCCATAAATTACATCTTCCCAAGGTGGAGATTTGCATGG 1656	
DB		
QY	1524 CCGGCTGTACAAAGAGCTACTCCATAAATTACATCTTCCCAAGGTGGAGATTTGCATGG 1583	
DB		
QY	1657 GAGACTCCGCTGACAAACCCCTTAAGCGCAATAATAGCTATATCTTCTATACCATGGCAA 1716	
DB		
QY	1584 GAGACTCCGCTGACAAACCCCTTAAGCGCAATAATAGCTATATCTTCTATACCATGGCAA 1643	
DB		
QY	1717 TATGTGGCATGCTCTGGATTCATTTCCTGTCGCAAGAGAGGTGAACAGAGGGCGGAAGAAA 1776	
DB		
QY	1644 TATGTGGCATGCTCTGGATTCATTTCCTGTCGCAAGAGAGGTGAACAGAGGGCGGAAGAAA 1703	
DB		
QY	1777 TGGAGAAGCTGACATGGGCTTAATCAGACTTCCAAAGAGCAATTCGAATGGACAGTTACA 1836	
DB		
QY	1704 TGGAGAAGCTGACATGGGCTTAATCAGACTTCCAAAGAGCAATTCGAATGGACAGTTACA 1763	
DB		
QY	1837 CCAGTTACTGCNAATGCTGTGTGACCTTCACTAGCATCTGAGATAGACATGAGTGCA 1896	
DB		
QY	1764 CCAGTTACTGCNAATGCTGTGTGACCTTCACTAGCATCTGAGATAGACATGAGTGCA 1823	
DB		

[illegible]

Db	2903	CCCTCCTGTCAGTAGTGGCAGGATCTATTGGCATAATTCGGGAGCTCTTTAGAGGGATGAG	2962
Qy	3037	GTTCCTTTGAACACAGTGAAAAATTTAAATTAGTAACATTTTTCGAAAGCAGGTTTATTGACTG	3096
Db	2963	GTTCCTTTGAACACAGTGAAAAATTTAAATTAGTAACATTTTTCGAAAGCAGGTTTATTGACTG	3022
Qy	3097	TTATTGCTAAGAAGAAGTAAGAAAGAAAAAGCCCTGTTGGCAATCTTGGTATTTCCTTTAA	3156
Db	3023	TTATTGCTAAGAAGAAGTAAGAAAGAAAAAGCCCTGTTGGCAATCTTGGTATTTCCTTTAA	3082
Qy	3157	GATTTCTGGCAGCTGTGGGATGGATGAATGAAGTGGATGTGAATCTTGGCAAGTTAAAT	3216
Db	3083	GATTTCTGGCAGCTGTGGGATGGATGAATGAAGTGGATGTGAATCTTGGCAAGTTAAAT	3142
Qy	3217	GGGACAGCCCTTCATGTTCAATTTGTCTTACCTCTTAACCTGAATAAAAAAGCCTACAGTTTT	3276
Db	3143	GGGACAGCCCTTCATGTTCAATTTGTCTTACCTCTTAACCTGAATAAAAAAGCCTACAGTTTT	3202
Qy	3277	TAGAAAAA 3285	
Db	3203	TAGAAAAA 3211	
RESULT 6			
HUMGLVRLX			
LOCUS	HUMGLVRLX	3220 bp	mRNA linear PRI 08-NOV-1994
DEFINITION	Human leukemia virus receptor 1 (GLVR1) mRNA, complete cds.		
ACCESSION	L20859		
VERSION	L20859.1	GI:306769	
KEYWORDS	leukemia virus receptor 1.		
SOURCE	Homo sapiens (tissue library: lambda HGR6, 7, and 16; Clontech #1020b) cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	O'Hara,B., Johann,S.V., Klinger,H.P., Blair,D.G., Rubinson,H., Dunn,K.J., Sasse,P., Vitek,S.M. and Robins,T.		
TITLE	Characterization of a human gene conferring sensitivity to infection by gibbon ape leukemia virus		
JOURNAL	Cell Growth Differ. 1 (3), 119-127 (1990)		
MEDLINE	91175479		
PUBMED	2078500		
FEATURES	Location/Qualifiers		
source	1..3220		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="2q11-q14"		
	/cell_line="HL60"		
	/tissue lib="lambda HGR6, 7, and 16; Clontech #1020b"		
	1..3220		
gene	/gene="GLVR1"		
	371..2410		
CDS	/gene="GLVR1"		
	/codon_start=1		
	/product="leukemia virus receptor 1"		
	/protein_id="AA352572.1"		
	/db_xref="GI:306770"		
	/db_xref="GB:G00-125-248"		
	/translations="MATLITSTTAATAAAGPLVDYLMMLILGIIAFVLAFSVGANDV ANSGFTAVGSGVTVLQACPLASIFETVGSVLLGAKVSETIRKGLIDVMYNSQTQGL MAGSVAMFGSAVQLVAFSLKPLIPSTGHCIVGATGFSLVAKGQGVKWSLEIKIWM SLKEDHEETLQVLDIENKHPVSEVGVPATVLPQAVVEERTVSPKLGDEAPERERLP SVLKEBTSIDSTVNGAQLPNNGLVQFSQAVNQINSSGHSQYHTVHKDSGLYKELL HKHLAKNVCMDGKQPLRRNSVTSYTMACGMPLDSFRAKEQEGKEEMKLTW PNADSKRIRBDSVTSYCNASVLDHSAEIDMSVKAAAGLDGRKNGSLSEEWVDDQ PEVSLVQFQQLLTACFGSAHGNDVSNAGPLVALYLVYDGTGVSSTKATPIWLL YGVGVICVGVGWRVRIQTMGKLTPIFTSPGFSIELASALTIVIASNLGLPISTH CKVGSVSVGWLRSKAXVDVFRNFRNFMFVTVPSIGVISAAIMAIFRIVILRM"		
BASE COUNT	783 a	763 g	964 t

[illegible]

Db 684 GTCAGTCAGTCATGTTTGGTTCTGCTGTGTGGCAACTCGTGGCTTCGTTTGTAAAGC 743
Qy 817 TCCCTATTTCTGGAACCCATTTGTTTGGTGGCAACTATTGGTTTCTCCCTCGTGGCAA 876
Db 744 TCCCTATTTCTGGAACCCATTTGTTTGGTGGCAACTATTGGTTTCTCCCTCGTGGCAA 803
Qy 877 AGGGCAGAGAGGGTGTCAAGTGGTCTGAACCTGATATAAAATTTGTATGTTGGTTGGTGT 936
Db 804 AGGGCAGAGAGGGTGTCAAGTGGTCTGAACCTGATATAAAATTTGTATGTTGGTTGGTGT 863
Qy 937 CCCCACTGCTTTCTGGAATTATGCTCTGGAATTTTATCTTCTGTTGGTCTGTGATTCATCC 996
Db 864 CCCCACTGCTTTCTGGAATTATGCTCTGGAATTTTATCTTCTGTTGGTCTGTGATTCATCC 923
Qy 997 TCCATAAGGCAGATCCAGTTCCTAATGTTTTCGAGCTTTTGCAGTTCCTATGCTGCA 1056
Db 924 TCCATAAGGCAGATCCAGTTCCTAATGTTTTCGAGCTTTTGCAGTTCCTATGCTGCA 983
Qy 1057 CAGTTGGGAATAAACCTCTTTTCCATCATGTATACCTGGAGACCGTTGCTGGCTTTTGACA 1116
Db 984 CAGTTGGGAATAAACCTCTTTTCCATCATGTATACCTGGAGACCGTTGCTGGCTTTTGACA 1043
Qy 1117 AACTTCTCTGTGGGTACCATCTCATCTCCGTGGATGTGCAGTTTCTGTGCCCTTA 1176
Db 1044 AACTTCTCTGTGGGTACCATCTCATCTCCGTGGATGTGCAGTTTCTGTGCCCTTA 1103
Qy 1177 TCGTCTGGTCTTTGTATGCTCCAGAGTGAAGAGAAAATTCGAAAGAGATAAAGTGTGA 1236
Db 1104 TCGTCTGGTCTTTGTATGCTCCAGAGTGAAGAGAAAATTCGAAAGAGATAAAGTGTGA 1163
Qy 1237 GTCTTCTGAAAGCCCTTAAATGGAAGAAAAGAAATAGCTTGAAGAGAACCATGAAGAAA 1296
Db 1164 GTCTTCTGAAAGCCCTTAAATGGAAGAAAAGAAATAGCTTGAAGAGAACCATGAAGAAA 1223
Qy 1297 CAAGTGTCTGTGGTGTATTTGAAACCAAGCATCTCTGTTTCTGAGGTAGGGCCCTGCA 1356
Db 1224 CAAGTGTCTGTGGTGTATTTGAAACCAAGCATCTCTGTTTCTGAGGTAGGGCCCTGCA 1283
Qy 1357 CTGTGCCCTCCAGCGCTGTGGTGGAGAGAGAAACAGTCTCATTTCAAACTTTGGAGATTTGG 1416
Db 1284 CTGTGCCCTCCAGCGCTGTGGTGGAGAGAGAAACAGTCTCATTTCAAACTTTGGAGATTTGG 1343
Qy 1417 AGGAAGTCCAGAGAGAGAGGCTTCCAGCGTGGACTTGAAGAGAGAAACCAAGCATAG 1476
Db 1344 AGGAAGTCCAGAGAGAGAGGCTTCCAGCGTGGACTTGAAGAGAGAAACCAAGCATAG 1403
Qy 1477 ATAGCACCGTGAATGGTGCAGTGCAGTTGCCCTTAATGGGAACTTGTCCAGTTCAGTCAAG 1536
Db 1404 ATAGCACCGTGAATGGTGCAGTGCAGTTGCCCTTAATGGGAACTTGTCCAGTTCAGTCAAG 1463
Qy 1537 CCGTCAGCAACCAAAATAAACTCCAGTGGCCACTACCAAGTATCACACCGTGCATAAGGATT 1596
Db 1464 CCGTCAGCAACCAAAATAAACTCCAGTGGCCACTCCAGTATCACACCGTGCATAAGGATT 1523
Qy 1597 CCGGCTGTACAAAGAGTACTCCATAAATTAACATCTTTCGCAAGGTGGAGATTGCATGG 1656
Db 1524 CCGGCTGTACAAAGAGTACTCCATAAATTAACATCTTTCGCAAGGTGGAGATTGCATGG 1583
Qy 1657 GAGACTCCGGTGACAAAACCTTTAAGCGGCAATAATAGCTATACTTCTATACCATGGCAA 1716
Db 1584 GAGACTCCGGTGACAAAACCTTTAAGCGGCAATAATAGCTATACTTCTATACCATGGCAA 1643
Qy 1717 TATGTGGCATGCTCTGGATTTCATTCGGTCCAAAGAGAGGTGAACAGAGGGCGAAGAAA 1776
Db 1644 TATGTGGCATGCTCTGGATTTCATTCGGTCCAAAGAGAGGTGAACAGAGGGCGAAGAAA 1703
Qy 1777 TGGAGAAGCTGACATGGCTTAATGAGACTCCAAAGAGCAATTCGAATGACAGCTTACA 1836
Db 1704 TGGAGAAGCTGACATGGCTTAATGAGACTCCAAAGAGCAATTCGAATGACAGCTTACA 1763
Qy 1837 CCAGTTACTGCAATGCTGTGCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA 1896

Db 1764 CCAGTTACTGCAATGCTGTGCTGTGACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA 1823
Qy 1897 AGGCAGAGATGGGTCTAGGTGACAGAAAGAAAGTAATAGGTCTCTAGAAAGTAATGATG 1956
Db 1824 AGGCAGAGATGGGTCTAGGTGACAGAAAGAAAGTAATAGGTCTCTAGAAAGTAATGATG 1883
Qy 1957 ACCAGGATAAGCTGAAAGTCT 2016
Db 1884 ACCAGGATAAGCTGAAAGTCT 1943
Qy 2017 TTGGGTCAATTCGCCCATGGTGGCAATGACGTAAAGCAATGCCATTTGGGGCTCTGGTTGCTT 2076
Db 1944 TTGGGTCAATTCGCCCATGGTGGCAATGACGTAAAGCAATGCCATTTGGGGCTCTGGTTGCTT 2003
Qy 2077 TATATTTGGTTTATGACACAGAGAGATGTTTCTTCAAAAAGTGGCAACACCAATATGGCTTC 2136
Db 2004 TATATTTGGTTTATGACACAGAGAGATGTTTCTTCAAAAAGTGGCAACACCAATATGGCTTC 2063
Qy 2137 TACTCTATGGTGGTGGTGGTATCTGTGTGGTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGT 2196
Db 2064 TACTCTATGGTGGTGGTGGTATCTGTGTGGTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGT 2123
Qy 2197 AGACCATGGGNAAGGATCTGACACCGATCACACCTCTAGTGGCTTCAGTATTGAAGTGG 2256
Db 2124 AGACCATGGGNAAGGATCTGACACCGATCACACCTCTAGTGGCTTCAGTATTGAAGTGG 2183
Qy 2257 CATCTGCCCTCACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2316
Db 2184 CATCTGCCCTCACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2243
Qy 2317 GTAAAGTGGGCTGT 2376
Db 2244 GTAAAGTGGGCTGT 2303
Qy 2377 GTCTCTTTCGTAAACATTTTATGCGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2436
Db 2304 GTCTCTTTCGTAAACATTTTATGCGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2363
Qy 2437 GTGTGCCCATCATGGCAATCTTCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2496
Db 2364 GTGTGCCCATCATGGCAATCTTCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2423
Qy 2497 TAAATTTGTGTCAATGTTTGGGACCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2556
Db 2424 TAAATTTGTGTCAATGTTTGGGACCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2483
Qy 2557 ACAGTGTAAACAGAGACTGACAGAGTCTTTTATTTGGGAGCCAGAGGAGGAGGAGTGT 2616
Db 2484 ACAGTGTAAACAGAGACTGACAGAGTCTTTTATTTGGGAG - CAGAGGAGGAGGAGTGT 2542
Qy 2617 TACTTGTCTATAACTGCTTTTGT 2676
Db 2543 TACTTGTCTATAACTGCTTTTGT 2602
Qy 2677 ATAGCCGGGTTCACATGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2736
Db 2603 ATAGCCGGGTTCACATGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2662
Qy 2737 GTACATATTTCTCTACTTTTGT 2796
Db 2663 GTACATATTTCTCTACTTTTGT 2722
Qy 2797 TGAAGATGACTTGTGATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2856
Db 2723 TGAAGATGACTTGTGATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2782
Qy 2857 CATGCTCTGCGT 2916
Db 2783 CATGCTCTGCGT 2842
Qy 2917 AAAAATAAATAACACTGACACTTCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2976
Db 2843 AAAAATAAATAACACTTCCCTTGT 2902

QY 770 TATGTTTGGTTCTGCTGTGTRGGCAACTCGRGCTTCGTTTTTGAAGCTCCCTATTTCTGG 829
DB 729 TATGTTTGGTTCTGCTGTGTRGGCAACTCGRGCTTCGTTTTTGAAGCTCCGATTTCTGG 788
QY 830 AACCCATTGATTTGTTGGTCAACTATTGGTTTCTCCCTCGTGGCAAGGGGAGAGGG 889
DB 789 GACCCATTGATTTGTTGGTCAACCACTTGGTTTCTCCCTTGTGGCAAGGGGAGAGGG 848
QY 890 TGTCAAGTGGTCTGAACATGATAAAAAATTGTGATGTCTTGGTTCTGTCGCCACTGCTTTC 949
DB 849 TGTCAAGTGGTCTGAACATGATAAAAAATTGTGATGTCTGTTGGTTCTCTCCGCTGCTTTC 908
QY 950 TGGAAATTATGTCGGAAATTTTATCTTCTCGGTTCTGTGCAATTCATCTCCATAAGGCAGA 1009
DB 909 TGGTATTATGTCGGAAATTTTATCTTCTCTGTTCTGTGCGGTTTCATCTCCGTAAGGCAGA 968
QY 1010 TCCAGTTCTTAATGGTTTCCGAGCTTTCGAGCTTTCAGGTTTCTATGCTCGCACAGTTGGATATA 1069
DB 969 TCCGGTTCTTAATGGCTTACGAGCTTTACCAATTTTTTATGCTGCGCAATCGGAATCAA 1028
QY 1070 CTTCTTTTCCATCATGTATCTAGGACACCGTTTCTGCTGGGCTTTGACAAAATCTTCTCTGTG 1129
DB 1029 CTTCTTTTCCATTTATGTATCTAGGACACCGTTTCTGCTGGGCTTTGACAAAATCTTCTCTGTG 1088
QY 1130 GGTGTACATCTCATCTCGGTGGGATGTGAGATTTTCTGTGCGCTTATCGTCTGTTCTT 1189
DB 1089 GGTGTACATCTCATCTCGGTGGGATGTGAGATTTTCTGTGCGCTTATCGTCTGTTCTT 1148
QY 1190 TGTATGCTCCAGGATCAAGAGAAATTTGAACGAGAAATAAGTGTAGTCTCTGAAAG 1249
DB 1149 TGTATGCTCCAGGATCAAGAGAAATTTGAACGAGAAATAAGTGTAGTCTCTGAAAG 1208
QY 1250 CCCCTTAATGGAAAAAGAAATAGCTTGAAGAGAACCATGAAGAAACAAAGTTGCTCT 1309
DB 1209 TCCCTTAATGGAAAAAGAGCAACTTAAAGAGAACCATGAAGAAACAAAGTGGCTCC 1268
QY 1310 TGGTGATATGAAAAACAAGCATCTCTGTTTCTGAGGTAGGCGCTTGCACCTGTGCCCCCA 1369
DB 1269 TGGAGACGTTGAGCATAGGAATCTCTGTCTGAGGTAGTGTGTGCCACTGGGCCACTCCG 1328
QY 1370 GGTGTGGTGGAGGAGAACAGTCTCATCTCAAACTTGGAGATTGGAGAGAGCTCCAGA 1429
DB 1329 GGTGTGGTGGAGGAGAACAGTGTCTTCAAACTTGGAGCTTGGAGAGGCTCCGGA 1388
QY 1430 GAGAGAGGCTTCCAGCGTGGACTTGAAGAGGAAACAGCATAGATAGCACCGTGAA 1489
DB 1389 GCGAGAGCGGCTTCCCA---TGGACCTGAAGGAGAGACAGCATAGACAGCACCATCAA 1445
QY 1490 TGGTGCAGTGCAGTTCCCTTAATGGGAACTTGTCCAGTTTCAAGTCAAGCCGTGAGCAACA 1549
DB 1446 TGGTGCAGTGCAGTTCCCTTAATGGGAACTTGTTCAGTTTCAAGTCAAGTGTGAGCAACA 1505
QY 1550 AATAAATCCAGTGGCCACTACAGTATCACACCGTGCATAGGATTCGGCGCTGTACAA 1609
DB 1506 GATCAACTCCAGTGGCCACTATCAGTATCACACCGTGCATAGGATTCGGCTGTGACAA 1565
QY 1610 AGAGCTACTCCATAAATATACATCTTCCCAAGTGGGAGATTGATCGGGAGACTCCGGTGA 1669
DB 1566 GGAGCTGCTCCATAGTTTACATCTGCGCAAGTGGGAGACTGCATGGGAGATTCTGGGA 1625
QY 1670 CAAACCTTAAGCGCAATAATAGCTATATCTTCTATACCATGGCAATATGTGGCATGCC 1729
DB 1626 CAAAGCCCTTGAGACGCAACAACAGCTACATCTCTACATATGGCAATATGTGGCATGCC 1685
QY 1730 TCTGGATTATCTCCGTGCCAAGAAGCTGAACGAAGGGCGGAAGAAATGGAGAGCTGAC 1789
DB 1686 CTTGGATTATCTCCGTGCCAAGAAGCTGAACGAAGGGCGGAAGAAATGGAGAGCTGAC 1745
QY 1790 ATGGCCTTAATGCAGACTCCAAGAGCGAATTCGAATGGAGAGTTACACCAAGTTACTGCAA 1849
DB 1746 ATGGCCTTAATGCAGATACCAAGAGCGGATTCGAATGGAGAGTTACACCAAGTTACTGCAA 1805

QY 1850 TCTGTGTCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGTCAAGGCAGAGATGGG 1909
DB 1806 TGCCTGTCTGACCTTCACTCCGAGTCTGAGATGAGACATGAGTGTGAAGGCTGGGATGGG 1865
QY 1910 TCTAGGTGACAAAAAGGAAAGTAATGGCTCTCTAGAAGAAATGGTATGACAGGATAAGCC 1969
DB 1866 CTTGGTGTGACAAAAAGGAAAGCAGTGGCTCTCTTGAAGAAATGGTATGACAGGATAAGCC 1925
QY 1970 TGAAGTCT 2029
DB 1926 CGAAGTGTCT 1985
QY 2030 CCATGGTGGCAATGACGTAAGCAATGCCCATTGGGCTCTGTTGCTTTATATATTTGGTTTA 2089
DB 1986 CCATGGTGGCAATGACGTCAGCAATGCCCATTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2045
QY 2090 TGACACAGGAGATGTTTCTTCAAAAGTGGCAACCAATATGGCTTCTACTCTATGGTGG 2149
DB 2046 T---AAAAAGAAAGCCTCTACAAAAGCGGCAACACCCATATGGCTTCTCTCTTTATGGTGG 2102
QY 2150 TGTGGTATCTGTGTGGTCTGTGGTGTGGGAAAGAGAGTTATCCAGACCATGGGAA 2209
DB 2103 TGTGGCATTTGCATGGGCTGTGGTGTGGGAAAGAGAGTTATCCAGACCATGGGAA 2162
QY 2210 GGATCTGACACCGATCACACCTCTAGTGGCTTTCAGTATTTGAACCTGGCATCTGCCCTCAC 2269
DB 2163 GGACCTGACCCCAATCACACCTCCAGTGGTTTCAGTATTTGAATGGCGCTCTGCCTTAAC 2222
QY 2270 TGTGTGATTGTCATCAATATTTGGGCTTCCCATCAGTACCAACACATTTGTAAGTGGGCTC 2329
DB 2223 TGTGTGATCGCATCAAAATTTGGGCTTCCCATCAGCAACACATTTGCAAAAGTGGGCTC 2282
QY 2330 TGTGTGTCTGTGTGGTCCGCTCCAGTCCAAAGAGCTGTGTGACTGGCGTCTCTCTTCGTA 2389
DB 2283 TGTGTGTCTGTGTGGTCCGATCAAGAAAGCTGTGTGACTGGGACTGTGTTTCGAAA 2342
QY 2390 CATTTTTATGGCTGGTTTGTACAGTCCCAATTTCTGAGTATTCAGTGTGCCATCAT 2449
DB 2343 CATTTTTATGGCTGGTTTGTACAGTCCCAATCTCTGGGTTATCAGTGTGCCATCAT 2402
QY 2450 GGCATCTTCAGATATGTCATCTCAGAAATGTGAAGCTGTTTGAGATTAATAATTTGTGTC 2509
DB 2403 GGCAGTATTCAGTATCATCTCTGCCAGTGTGACGC---TGGGGTCCAAAAGCTGTGTC 2458
QY 2510 AATGTTTGGGACCATCTTAGGTATTTCTGCTCCCTGCAAGAAATGATACAGTGTAAACAG 2569
DB 2459 AGTGTCTGGACCATTTGACATTTCTTG---TTCCTAGGAAACGCTCACAGTGT---GCTG 2516
QY 2570 AAGACTGACAAGAGTCTTTTATTTTGGGAGCCAGAGGAGGAAAGTGTACTTTGTGCTATA 2629
DB 2517 AAGACAGCAAGGGTCT---TAAAGGAGCGCTGGGAAGGAGTGAATTTTACACTATA 2571
QY 2630 ACTGCTTTTGTGCTAAATATGAAATTTGCTCAAAATTTAGCTGTGTAATAATAGCCCGGTTTC 2689
DB 2572 ATTGCTTTTGTGCTAAATATGACTATCTCAAAATTTAGCTATGTAATAATAGCCAGGTTTC 2631
QY 2690 CACTGGCTCTGCTCAGGTCCCTTTCTTCTGGGCTGTGAATTCCTGTACATATTTCTC 2749
DB 2632 CATTAATTCATTCGAAGT---CCCTTTTCTCTGGGCTATGAATTCCTGTACATATTTCTC 2690
QY 2750 TACTTTTTTGTATCAGGCTTCAATTCATATATGTTTTTAAATGTTGTCTCTGGAAGATCACTTG 2809
DB 2691 TAC---TTTTGTATCAGGCTCAATTCAGTATGTTTTTAAATGTTGTCTGTGAGATAACTTAG 2749
QY 2810 TGAATTTTTTCTTTTTTTTTTAAACCATGAAGAGCGGTTTGACAGAGCATGCTCTGCGTT 2869
DB 2750 -----GTGGGTTCTTTTTTAAACAGCCAGCCAGAGCCATTG---ATGGCAGCTACTGCTTT 2802
QY 2870 GTTGTGTTTACCAGCTCTGCGCTCACATGCACAGGATTTTAAACAAACAAATAAATACTA 2929
DB 2803 GTGCGCTTACCAGCTCTTCTCCCAACATGCACAGGATTTTAAACAAACATGTAACAGC 2862
QY 2930 CAACTTCCCTTGTAGTCTCTTATATAAAGTAGAGTCTTGGTACTCTGCGCTCTCTGTCAGT 2989

QY 1523 CCAGTTTCAGTCAAGCCGTGAGCAACCAAAATAAATCACTCCAGTGGCCACTACAGATATCACAC 1582
DB 1498 TCAGTTTCAGTCAAACTGTGAGCAACCAAGATCAACTCCAGTGGCCACTATCAGATATCACAC 1557
QY 1583 CGTGCATAAGGATTCGGGCTGTACAAAGAGCTACTCCATAAATACATCTTGGCAAGGT 1642
DB 1558 CGTGCATAAGGATTCGGGCTGTACAAAGAGCTACTCCATAAATACATCTTGGCAAGGT 1617
QY 1643 GGGAGATTCATGGGAGACTCCGGTGACAAACCCCTTAAGGCGCAATAAATAGCTATACTTC 1702
DB 1618 GGGAGACTCATGGGAGATTCGGGAGCAAGCCCTTGAGCGCAACACAGCTACACTTC 1677
QY 1703 CTATACCATGGCAATATGTGGCATGCTCTGGAATCATTTCCGTGCCAAAGAGGTGAACA 1762
DB 1678 CTACACTATGGCAATATGTGGCATGCTCCCTGGATTCATTCGGTGCCAAAGAGGTGAACA 1737
QY 1763 GAAGGCGAAGAAATGGGAAGCTGACATGGGCTTAATGAGACTCCAGAAAGGATTCG 1822
DB 1738 AAGGGAGATGAATGGAGACGCTGACATGGGCTTAATGAGACTCCAGAAAGGATTCG 1797
QY 1823 AATGACAGTTACACAGTTACTGCAATGCTGTGCTGACCTTCACTCAGCATCTGAGAT 1882
DB 1798 AATGACAGTTACACAGTTACTGCAATGCTGTGCTGACCTTCACTCAGCATCTGAGAT 1857
QY 1883 AGACATGAGTGTCAAGGCAAGATGGGTCTAGGTGACAGAAAGGAAGTAAATGGCTCTCT 1942
DB 1858 GGACATGAGTGTCAAGGCTGAGATGGGCTGGGTGACAGAAAGGAAGTAAATGGCTCTCT 1917
QY 1943 AGAAGATGATATGACCAAGTAAGCTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2002
DB 1918 TGAAGATGATATGACCAAGTAAGCTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1977
QY 2003 CTTTACAGCTGCTTTGGGTCATTCGCCCATGGTGGCAATGACGTAAGCAATGCCATTGG 2062
DB 1978 CTTTACAGCTGCTTTGGGTCATTCGCCCATGGTGGCAATGACGTAAGCAATGCCATTGG 2037
QY 2063 GCCTCTGGTGTCTTATATTTGGTTTATGACAGAGAGATGTTCTTCAAAAAGTGGCAAC 2122
DB 2038 CCCTCTGGTGTCTTATATTTGGTTTATGACAGAGAGATGTTCTTCAAAAAGTGGCAAC 2094
QY 2123 ACCAATATGGCTTCTACTCTATGTTGGTGTGGTATCTGTGTGCTGTGGGTTTGGG 2182
DB 2095 ACCAATATGGCTTCTACTCTATGTTGGTGTGGTATCTGTGTGCTGTGGGTTTGGG 2154
QY 2183 AAGAAGAGTTATCCAGACCATGGGAAGGATCTGACACCGATCAACACCTCTAGTGGCTT 2242
DB 2155 AAGAAGAGTTATCCAGACCATGGGAAGGATCTGACACCGATCAACACCTCTAGTGGCTT 2214
QY 2243 CAGTATTTGAATGGCATTCGCCCTCACTGTGTGATTTGCATCAAAATATTTGGCTTCCCAT 2302
DB 2215 CAGTATTTGAATGGCATTCGCCCTCACTGTGTGATTTGCATCAAAATATTTGGCTTCCCAT 2274
QY 2303 CAGTACACACATTTGAAGTGGGCTCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTG 2362
DB 2275 CAGCAACACATTTGAAGTGGGCTCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTG 2334
QY 2363 GCCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTG 2422
DB 2335 GCCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTG 2394
QY 2423 TTCTGGAGTTATCAGTGTGCTCATGTCGCAATCTTCAGATATGTCCTCAGATGTG 2482
DB 2395 CTCTGGGTTATCAGTGTGCTCATGTCGCAATCTTCAGATATGTCCTCAGATGTG 2454
QY 2483 AGCTCTTTGAGATTAATTTGTCATATGTTGGGACCATCTTAGGTATTCCTGCTCC 2542
DB 2455 ACCTCTTTGAGATTAATTTGTCATATGTTGGGACCATCTTAGGTATTCCTGCTCC 2509
QY 2543 CCGTGAAGATGATTAAGTGTGTTAAACAGAGACTGACAGAGTCTTTTATTTGGGAGCCA 2602
DB 2510 CTAGAGAAAGCTCACAGTGTGTTGCTGAAGACAGGCAAGGGTCTCTTAAAGGAGCGG 2563

QY 2603 GAGGAGGAAGTGTACTTGTGCTATAAATGCTTTTGTGCTAAATATGAATTTGCTCAAA 2662
DB 2564 TGGGAGGAAGTGTAAATTTACACTATAAATGCTTTTGTGCTAAATATGAATTTGCTCAAA 2623
QY 2663 ATTAGCTGTGTAATAATAGCCGGTTCACATGGCTCCTGCTGAGCTGCCCTTCTCTCTG 2722
DB 2624 ATTAGCTGTGTAATAATAGCCGGTTCACATGGCTTCCATGATTCATTTCCAAAGT-CCCTTTTCTCTCTG 2682
QY 2723 GGCTGTGAATTCCTGTACATATTTCTACTTTTGTATCAGGCTTCAATTTCAATTTATGCT 2782
DB 2693 GGTATGAATTCCTGTACATATTTCTCTAC-TTTTGTATCAGGCTCAATTTCCAGTATGT 2741
QY 2783 TTTAATGTGTCTCTGAAGATGACTTTGTGATTTTCTTTTCTTTTAAACCATGAAGA 2842
DB 2742 TTTAATGTGTCTGTGA-----GATAAATTTAGTGGTGTCTTTTAAACAGCCAGCA 2793
QY 2843 GCCGTTTTCACAGAGCATGCTCTGCGTGTGTGTTTTCACAGCTTCTGCCCTCACATGCAC 2902
DB 2794 GAGCCATTTGATGGCATGTACTGCTTTGTGGGCTTCACAGCTTCTTCCCCAACATGCAC 2853
QY 2903 AGGGAATTTAAACAAACAAATAATACTACAACCTTCCCTTGTAGTCTCTTATATAAGTAGAG 2962
DB 2854 AGGGAATTTAAACAAACATGTAACCTGAAGCTTCCCTCCTCATAGTCTCTCATAGAAATAGTC 2913
QY 2963 TCCTTGTGTACTCTGCTCCTCTGTAGTGTGAGGATCTATTTGGCATATTCGGGAGCTT 3022
DB 2914 AC---GGCAGCTCTGCTC-CCTGTCTACTAGTGGCAGCTTCTGTG--ATGTGTGACAACTT 2967
QY 3023 CTTAGAGGATGAGTCTTTTGAACACAGTGAATAATTTAAAT---TAGTAACTTTTTCG 3079
DB 2968 CTTAGAGGCGGAGAAATCTTTG-GCACAGTGGAAATATAAGTTTCTAGTAACTCTTTTCG 3026
QY 3080 AAGCAGTTTATGACTGTATTGCTTAAGAAGAAAGTAAAGAAAGAAAGGCTGTGGCAAT 3139
DB 3027 AAACAGTTTCAAGGAC--ATGTGTCTAAGACAGCAGGAGACAAA-----GCCCTCGCGGT 3079
QY 3140 CTTGTTTATTTCTTTAAGATTTCTGGCAGTGTGGATGGATGAATGAAGTGAATGTGAA 3199
DB 3080 TGTGTTTATTTCTCTGAGATTTCTGGCAGTGTGGATGGTGAATGAAGTGAATGTGAA 3139
QY 3200 CTTTGGGCAAGTAAATGGGACAGCTTCCATGTTTCAATTTGCTCTACCTCTTAACTGAATA 3259
DB 3140 CTTTGGGCAAAATTTCAATGGGACAGCTTCCATGTTTCAATTTGCTCTACCTCTTAACTGAATA 3198
QY 3260 AAAAGCCTACAGTGTTTTGAAGAAAAAAA 3288
DB 3199 AAAAGCCTACAGTGTTTTGAAGAAAAAAA 3227

RESULT 10
MUSGLVR1
LOCUS MusGlvr-1 3227 bp mRNA linear ROD 25-AUG-1995
DEFINITION Murine Glvr-1 mRNA, complete cds.
ACCESSION M73696
VERSION M73696.1 GI:957211
KEYWORDS
SOURCE mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1. (bases 1 to 3227)
O'Hara, B.
TITLE /organism="Mus musculus"
/db_xref="taxon:10090"
/issue_type="thymus"
JOURNAL 1. 3227
COMMENT On Aug 25, 1995 this sequence version replaced gi:193553.
FEATURES
Location/Qualifiers
source
gene

CDS

```

410. .2455
/gene="Glvr-1"
/codon_start=1
/product="Glvr-1"
/protein_id="AAA74887.1"
/db_xref="GI:193554"
/translation="NESTVATITSTLAAVTASAPPKYDNLWMLILGFIIAFVLAFSVG
ANDVANSRGTVAGSVVTLKQACILASIPETVGSALLGAKVSETIRNGLIDVELNYET
QDLLWAGSVAMFGSAWOLVASFLKIPISGICVIGATIGFSLVANGOKGVKWBELI
KIVMGFVPLSLGMSGLIFLVRFAFLKADVPNGRLRPIFYACTIGINLSIM
YTGAPLGLDPLKPLIILISVGCALIVWFVCPMKKIKIREVKSPSESPLM
EKKSLKDEHETKMAPGDVHRNPVSEVVCATGPLRAVVEBRTVSKLGLDEEAPLR
ERLNDLKEETSIDTINGAQLPQNGNLVQFSQTVSNQINSSGHVQYVHVHDKSLYK
ELUHLKHLAUGCGDGDGDKPELRNNSYSTIMAIICMPLDSFRKGEQKQDBMET
LWPNADTKRIRMDISYTSYCNVADLHSEMDMSVKAEMGLGDRKQSSGSELEWYI
ODKPEVLSLQILITACFGFAHGNDVSNALGPLVALYLVYKQEAATKAATLWIL
LLYGGVIGCMGLWVWRRIQTMGKDLPTIPSSGFSIELASALTAVTASINIGLPIST
THCKVSVSVSGWLRKSKAVDWRLFRNIFMAWFTVPIPSGVISAAIMAVFYIILPV"
BASE COUNT      781 a   740 c   802 g   904 t
ORIGIN

```

```

Query Match      64.8%; Score 2133; DB 10; Length 3227;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2662; Conservative 0; Mismatches 420; Indels 67; Gaps 21;

QY 162 CCCTCCCTTTCCCTGGATGAACCTGCGTCCCTTCTCTCTCGCCCATGGAATTCGCTC 221
DB 124 CACTTCGTCCTCCCGAGATGAACCTGCGTCCCTTCTCTCTCGCCCATGGAATTCGCTC 183

QY 222 CGTGCTTTTGGCCCTCCTGAGCCAAAGAAACCCAGACAGATGCCCATAC---GCAG 278
DB 184 CGTGCTTTTGGCCCTCAGAGCCAAAGAAACCCAGACAGATGCCCATGAGCAGCAG 243

QY 279 CGTATAGCAGTAACATCCCGAGCTCGGTTCTGTGCGGTAGTTTACAGTATTTAATTTTAT 338
DB 244 CGTATAGCAGTAACATCCCGAGCTCGGTTCTGTGCGGTAGTTTACAGTATTTAATTTTAT 303

QY 339 ATAAATATATATTTATTTATAGCATTTTGTATACCTCATATCTGTTTACACATCTGA 398
DB 304 ATAATATATATATTTATTTATAGCATTTT---GATACCTCATATCTGTTTACACATCTCAA 360

QY 399 AAGCGCTCAGTAGTTCTCTTA---CTAAACCACTACTCCAGAGAA-----444
DB 361 AAGCGCTTAGTAATTTCTTATTATTTAAGAACCACTACACTAGAGATGGAATCTAC 420

QY 445 --TGCAACCGGTGATTTACAGTACTACAGTCTACCGCGCTTCTGGTCTTTGGTGA 502
DB 421 TGTGGCAACGATTAAGTACCTAGTCTGCTGTTACTGCTTCCGCTCCACCGAAGTATGA 480

QY 503 CTACCTATGATGCTCATCTGGGCTTATTATTTGATTTGCTTGGCATTTCTCGTGGG 562
DB 481 CAATCTATGATGCTCATCTGGGCTTTCATCTATTTGCTTGGCATTTCTCGTGGG 540

QY 563 AGCAATATGATAGCAATTTCTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCCCTGAA 622
DB 541 AGCAATATGATAGCAATTTCTTTGGTACAGCTGTGGGCTGAGGTGTAGTGACCCCTGAA 600

QY 623 GCAAGCCTGCATCCTAGCTAGCATCTTTGAAACAGTGGGCTGTCTTTACTGGGGGCCAA 682
DB 601 GCAAGCCTGCATCCTAGCTAGCATCTTTGAAACAGTGGGCTGTCTTTACTGGGGGCCAA 660

QY 683 AGTGAGGAACCACTCGGNAGGCTTGATTGACGTGGAGATGTACAACTCGACTCAAGG 742
DB 661 AGTGAGGAACCACTCGGNAGGCTTGATTGACGTGGAGATGTACAACTCGACTCAAGA 720

QY 743 GCTACTGATGGCGGCTCAGTCACTGCTATGTTGGTCTGTGCTGTGGCAACTCGTGGC 802
DB 721 TCTGCTCATGGCTGGCTCCGTAGTCTGCTATGTTGGTCTGTGCTGTGGCAGCTCGTGGC 780

QY 803 TTGGTTTTGAAGCTCCCTATTTCTGGAACCCCAATTTGATTTGTTGGTGGCAACTATTTGGTTT 862
DB 781 TTGGTTTTGAAGCTCCGATTTCTGGAGCCCATTTGATTTGTCGGTGCAACCATTTGGTTT 840

```

```

QY 863 CTCCCTCGTGCAAAAGGGCAGAGGGTGTCAAGTGGTCTGAACTGATAAAATTTGTGAT 922
DB 841 CTCCCTGTGTGCAAAATGGCAGAGGGTGTCAAGTGGTCTGAACTGATAAAATTTGTGAT 900

QY 923 GTCTTGGTTCTGTGCCCACTGCTTTCTGGAATTAATGTCTGGAATTTTATTTCTCTCGT 982
DB 901 GTCTGGTGTCTGTCTCTCGCTGCTTTCTGGTATTATGTCTGGAATTTTATTTCTCTCTGT 960

QY 983 TCGTGCAATTCATCTCCATAAGGCAGATCCAGTTCCTTAATGGTTTGCAGCTTTGCCAGT 1042
DB 961 TCGTGCGTTTCATCTCCGTCCGTAAAGGCAGATCCCGTTCCTAATGGCTTTACAGCTTTACCAAT 1020

QY 1043 TTTCTATCCCTGACAGTGTGGAATAAACCCTCTTTTCCATCATGTATATCTAGGACACCGTT 1102
DB 1021 TTTTATGCCCTGACCAATCGGAATCAACCTCTTTTCCATTAATGTATATCTAGGACACCGTT 1080

QY 1103 GCTGGGCTTTGACAAACCTTCTCTGTGGGTACCATCTCATCTCGTGGGATGTGCAAT 1162
DB 1081 GCTGGGCTTTGACAAACCTTCTCTGTGGGTACCATCTCATCTCGTGGGATGTGCAAT 1140

QY 1163 TTTCTGTGCCCTTATCTGCTCTGTTCTTTGTATGTGCCAGGATGAGAGAAAATTTGAACG 1222
DB 1141 TTTCTGTGCCCTTATCTGCTCTGTTCTTTGTATGTGCCAGGATGAGAGAAAATTTGAACG 1200

QY 1223 AGAAATAAAGTGTAGTCTCTCTGAAAGCCCTTTAATGGAATAAAGAAATAGCTTCAAGA 1282
DB 1201 AGAAGTAAAGTGTAGTCTCTGAAAGTCCCTTAATGGAATAAAGAAATAGCTTCAAGA 1260

QY 1283 AGACCATGAAGAAACAAAGTTGTCTGTTGGTGATATTGAAACAAAGCAATCTGTTTCTGA 1342
DB 1261 AGACCATGAAGAAACAAAGATGGCTCTCTGGAGACGTTGAGCATAGGAATCTCTGTCTGA 1320

QY 1343 GGTAGGCTTGCCTGCTGCTCCAGCTGCTGGTGAGGAGAGAAACAGTCTCATTTCAA 1402
DB 1321 GGTAGTGTGTGCCACTTGGGCCACTCCGGGCTGTGGTGAGGAGAGACGGTGTCTCATTTCAA 1380

QY 1403 ACTTCGAGATTTGGAGGAAGCTCCAGAGAGAGAGAGAGCTTCCACCGCTGGAGCTTCAAGA 1462
DB 1381 ACTTGTGTGACCTGGAGAGGCTCCGGAGGAGAGAGCGGCTTCCCA---TGGACCTGAGGA 1437

QY 1463 GGAAACCAAGCATAGATAGACCGTGAATGGTGTGAGTGTGAGTGTGCTTAATGGGAACTTGT 1522
DB 1438 GGAGACCAAGCATAGACAGCACCATCAATGGTGTGAGTGTGAGTGTGCTTAATGGGAACTTGT 1497

QY 1523 CCAGTTCAGTCAAGCGCTCAGCAACCAATAACTCCAGTGGCCACTACCAAGTATCACAC 1582
DB 1498 TCAGTTCAGTCAAACTGTGAGCAACAGATCAACTCCAGTGGCCACTATCAGTATCACAC 1557

QY 1583 CGTGCAATAGGATTTCCGGCTGTACAAAGAGCTACTCCATATAATTTACATCTTGGCCAAAGT 1642
DB 1558 CGTGCAAAAGATTTCTGGCTGTACAAAGAGCTGTCTCCATTAAGTTTACATCTGGCCAAAGT 1617

QY 1643 GGGAGATTCATGGGAGACTCCGGTGAACAAACCTTTAAGGCGCAATAATAGCTATCTTTC 1702
DB 1618 GGGAGACTGATGGGAGATTTCTGGGGAACAAGCCCTTGAGACGCAACAACAGCTACACTTC 1677

QY 1703 CTATACCATGGCAATATGTGGCATCGCTTGGATTCATTTCCGTGCCAAGAGAGTTGAACA 1762
DB 1678 CTACACTATGGCAATATGTGGCATGCCCTGGATTCATTTCCGTGCCAAGAGAGTTGAACA 1737

QY 1763 GAAGGCGCAAGAAATTTGGAGAACTGACATGGCTAATGAGAGCTCCAAAGAACGAATTCG 1822
DB 1738 AAGGGAGATGAATTTGGAGACGCTGCATGGCTAATGCAGATACCAAGAGCGGATTCG 1797

QY 1823 AATGACAGTTTACACAGTACTGCAATGTCTGTGTGACCTTCTACTCAGCATCTGAGAT 1882
DB 1798 AATGACAGTTTACACAGTACTGCAATGTCTGTGTGACCTTCTACTCAGCATCTGAGAT 1857

QY 1883 AGACATGATGTCAAGGCAGAGATGGTCTAGTGTGACAAAGAGAGTAAGTGGCTCTCT 1942
DB 1858 GACATGATGTGAAAGGCTGAGATGGGCTGGGTGACAGAAAGAGAGTGGCTCTCTCT 1917

QY 1943 AGAAGAAATGGTATGACCAAGATAAGCCTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCT 2002

```



```
Db 1918 TGAAGAAATGGTATGACGAGTAAGCCTGAAGTGTCCCTTCTCTCCAGTTCCTGAGAT 1977
Qy 2003 CTTTACAGCTGCTCTTTGGGTCAATTCGCCCATGGTGGCAATGACGTAAGCAATGCCATTGG 2062
Db 1978 CTTTACAGCTGCTCTTTGGGTCAATTCGCCCATGGTGGCAATGACGTAAGCAATGCCATTGG 2037
Qy 2063 GCCTCTGGTGTCTTTATATTTGGTATTGACACAGAGATGTTTCTTCAAAAGTGCGCAAC 2122
Db 2038 CCCTCTGGTGTCTTTATATTTGGTATTGACACAGAGATGTTTCTTCAAAAGTGCGCAAC 2094
Qy 2123 ACCAATATGGCTCTTCTACTCTATGTGGTGTCTGATCTGTGTTGTTGTTGTTGTTGTTGTTG 2182
Db 2095 ACCAATATGGCTCTTCTACTCTATGTGGTGTCTGATCTGTGTTGTTGTTGTTGTTGTTGTTG 2154
Qy 2183 AAGAAGAGTTATCCAGACCATGGGAAGGATCTGCACACCGATCAGACCCCTCTAGTGGCTT 2242
Db 2155 AAGAAGAGTTATCCAGACCATGGGAAGGATCTGCACACCGATCAGACCCCTCTAGTGGCTT 2214
Qy 2243 CAGTATTGAACTGGGATCTGCCCTCACTGTGTGTGATGTCATCAAAATATTTGGCTTCCCAT 2302
Db 2215 CAGTATTGAACTGGGATCTGCCCTCACTGTGTGTGATGTCATCAAAATATTTGGCTTCCCAT 2274
Qy 2303 CAGTACAACACATTTGAAGTGGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2362
Db 2275 CAGTACAACACATTTGAAGTGGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2334
Qy 2363 GCCTGTGACTGGCTCTCTTTGTAACATTTTATGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2422
Db 2335 GCCTGTGACTGGCTCTCTTTGTAACATTTTATGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2394
Qy 2423 TTCTGGAGTTATCAGTGTGCGCATCATGGCAATCTTCAGATATGTCATCTCAGAAATGTG 2482
Db 2395 CTCTGGGTTATCAGTGTGCGCATCATGGCAATCTTCAGATATGTCATCTCAGAAATGTG 2454
Qy 2483 AGCTGTGTTGAGATTAATAATTTGTTCAATGTTTGGGACCATCTTAGGTATTTCTGCTCC 2542
Db 2455 AGCTGTGTTGAGATTAATAATTTGTTCAATGTTTGGGACCATCTTAGGTATTTCTGCTCC 2509
Qy 2543 CCTGAAGAATGATTAAGTGTGTTAAGCAAGACGCTGACAGAGTCTTTTATTTGGGAGCCA 2602
Db 2510 CTAGAGAAACGCTCAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2563
Qy 2603 GAGGAGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2662
Db 2564 TGGGAAGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2623
Qy 2663 ATTAGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2722
Db 2624 ATTAGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2682
Qy 2723 GGCTGTGAATTCCTGTACATATTTCTTACTTTTGTATGAGGCTTCAATTCATTTATGT 2782
Db 2683 GGCTGTGAATTCCTGTACATATTTCTTACTTTTGTATGAGGCTTCAATTCATTTATGT 2741
Qy 2783 TTTAATGTTGTTCTGTAAGATGACTGTGTAATTTTTTTTTTTTTTTTTTAAACCATGAAGA 2842
Db 2742 TTTAATGTTGTTCTGTA -----GATAACTTTAGTGGGTTCTTTTTTAAACAGCCAGCA 2793
Qy 2843 GCCGTTTGACAGAGATGCTGTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2902
Db 2794 GAGCCATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2853
Qy 2903 AGGGATTTTAAACAAACAAATATAACTTACAACTTCCCTTCTGTTGTTGTTGTTGTTGTTGTTG 2962
Db 2854 AGGGATTTTAAACAAACAAATATAACTTACAACTTCCCTTCTGTTGTTGTTGTTGTTGTTGTTG 2913
Qy 2963 TCCCTTGGTACTCTGCCCTCTCTGTCAGTAGTGGCAGATCTATTGGCATAATTCGGGAGCTT 3022
Db 2914 AC---GGCACTCTGCTC-CCTGTCACTAGTGGCAGGTTCTGTTG--ATGTGTGACAACTT 2967
Qy 3023 CTTAGAGGATGAGTGTCTTTTGAACACAGTGAAATTTAAAT---TAGTAACTTTTGTG 3079
```

```
Db 2968 CTTAGAGGCGCCGAGAAATCTTTG-GCACAGTGGAAATATAAGTTTGTAGTAACTCTTTTGC 3026
Qy 3080 AAGCAGTTTATTGACTGTATTGCTTAAGAAGAAAGTAAAGAAAGAAAGCTGTGTGGCAAT 3139
Db 3027 AAACAGTTTCAAGGAC-ATGTTGCTTAAGAGCAGGAGACAAA-----GCCCTGGCGGT 3079
Qy 3140 CTTGCTTATTTCTTTTAAAGATTTCTGCGCAGTGTGGGATGATGAATGAAGTGAATGTGAA 3199
Db 3080 TGTGCTTATTTCTTCTGAGATTTCTGCGCAGTGTGGGATGGTGAATGAAGTGAATGTGAA 3139
Qy 3200 CTTTGGGCAAGTTAAATGGGACAGCTTCCATGTTTCAATTTGCTACCTCTTAACTGAATA 3259
Db 3140 CTTTGGGCAAAATCAATGGGACAGCTTCCATGTTTCACTGTCTACCTCTTAACTGAATA 3198
Qy 3260 AAAAAGCCTACAGATTTTAAAGAAAAAAA 3288
Db 3199 AAAAAGCCTACAGATTTTAAAGAAAAAAA 3227

RESULT 11
LOCUS I11795 3260 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 3 from Patent US 5414076.
ACCESSION I11795
VERSION I11795.1 GI:909739
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3260)
AUTHORS O'Hara,B.M.
TITLE DNA encoding gibbon ape leukemia virus receptor
JOURNAL Patent: US 5414076-A 3 09-MAY-1995;
FEATURES
source 1..3260
BASE COUNT 790 a 747 c 809 g 914 t
ORIGIN

Query Match 64.8%; Score 2133; DB 6; Length 3260;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2662; Conservative 0; Mismatches 420; Indels 67; Gaps 21;

Qy 162 CCCTCCCTTTTCCCTGGATGAACCTTGGCTGCTTCTTCTCTCCGCGCATGGAAATTCGTCTC 221
Db 157 CACTTCGTCCTCCCGCAGGATGAACCTTGGCTGCTTCTTCTCTAAATCCGCGCATGGAAATTCGTCTC 216
Qy 222 CGTGCTTTTAGCCTCTCTGAGCCAAAGAAACCCAGACAAACAGATGCCCATAC---GCAG 278
Db 217 CGTGCTTTTAGCCTCTCTGAGCCAAAGAAACCCAGACAAACAGATGCCCATAC---GCAG 276
Qy 279 CGTATAGCAGTAACCTCCCGCAGCTCGGTTTCTGTCGCTAGTTTACAGTATTTTAAATTTAT 338
Db 277 CGTATAGCAGTAACCTCCCGCAGCTCGGTTTCTGTCGCTAGTTTACAGTATTTTAAATTTAT 336
Qy 339 ATAATATATATTATTATTATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTTGA 398
Db 337 ATAATATATATTATTATTATAGCATTTT---GATACCTCATTCGTTTACACATCTCAA 393
Qy 399 AAGCGCTCAGTAGTCTCTTA---CTAAACAACCACTACTCCAGAGAA-----444
Db 394 AAGCGCTTAGTAATTTCTTATTATTAAAGAACCACTACTAGAGAAATGGAATCTAC 453
Qy 445 ---TGGCAACGGTATTACCACTACTAGTCTACCGCGCTTCTGGTCTTCTTGGTGA 502
Db 454 TGTGCAACGATTTACTAGTACCTTAGCTGTGTTACTGCTTCCGCTCCACCGAAGTATGA 513
Qy 503 CTACCTATGGATGCTCATCTCGGGCTTCATTTATTCATTTGCTTTGGGATTTCTCGGTGG 562
Db 514 CAATCTATGGATGCTCATCTCGGGCTTCATTCATTCATTCATTTGCTTTGGGATTTCTCGGTGG 573
Qy 563 AGCCAAATGATGAGCAAAATCTTTTGGTACAGCTGTGGCTCAGGTGTAGTGAACCTGAA 622
```

Db 574 AGCCAATGATGATAGCAAAATTCGTTGGTACAGCTGTAGGCTCAGGTGTAGTACGCCCTGAA 633
Qy 623 GCAAGCCTGCATCTAGCTAGCATCTTTGAAAACAGTGGGCTCTGTCTTACTGGGGCCAA 682
Db 634 GCAAGCCTGCATCTTAGCTAGCATCTTCGAAAACAGTGGGCTCGGCTTCTGGGGCCAA 693
Qy 683 AGTGAGCGAAACATCCGGAAGGCTTGATGACGTGGAGATGTACAACTCGACTCAAGG 742
Db 694 AGTGAGCGAAACATCCGGAAGGCTTGATGACGTGGAGATGTACAACTCGACTCAAGG 753
Qy 743 GCTACTGATGGCGGCTCAGTCAGTCTATGTTGGTCTGTCTGTGTGGCAACTCGTGGC 802
Db 754 TCTGCTCATGGCTGGCTCGCTCAGTCTATGTTGGTCTGTGTGTGGCAACTCGTGGC 813
Qy 803 TTCGTTTTGAGCTCCCTATTTCTGGAACCCATTGTATGTTGGTGGCAACTATTGGTTT 862
Db 814 TTCGTTTTGAGCTCCCTATTTCTGGAACCCATTGTATGTTGGTGGCAACTATTGGTTT 873
Qy 863 CTCCCTCGTGGCAAGGGGCGAGGGGTGTCAAGTGGTCTGAACTGATAAAAAATTGTAT 922
Db 874 CTCCCTGTGGCAATGGGCGAGAGGGGTGTCAAGTGGTCTGAACTGATAAAAAATTGTAT 933
Qy 923 GTCTGGTGTGTCGCCACTGCTTTCTGGAATTTATGTCTGGAATTTATTTCTCTGTGT 982
Db 934 GTCTGGTGTGTCGCCACTGCTTTCTGGAATTTATGTCTGGAATTTATTTCTCTGTGT 993
Qy 983 TCGTGCAATTCATCCTCCATAAGGCAGATCCAGTTCCTAATGTTTGGGAGCTTGGCAGT 1042
Db 994 TCGTGCAATTCATCCTCCATAAGGCAGATCCGTTCCATAATGTTTACGAGCTTTACCAAT 1053
Qy 1043 TTTCTATGCTGCACAGTTGGAAATAAACCCTTTTCCATCATGTATATCTGGAGCACCGTT 1102
Db 1054 TTTTATGCTGCACAAATCGGAATCAACCTCTTTTCCATTTATGTATCTGGAGCACCGTT 1113
Qy 1103 GCTGGCTTTGCAAACTTCTCTGTGGGTACCACTCTCATCTCGGTGGGATGTGCACT 1162
Db 1114 GCTGGCTTTGCAAACTTCTCTGTGGGTACCACTCTCATCTCGGTGGGATGTGCACT 1173
Qy 1163 TTTCTGTGCTTATCGTCTGGTCTTTGTATGTCAGGATGAAGAGAAAAATTGAAAGC 1222
Db 1174 TTTCTGTGCTTATCGTCTGGTCTTTGTATGTCAGGATGAAGAGAAAAATTGAAAGC 1233
Qy 1223 AGAAATAAAGTGTAGTCTCTGAAAGCCCTTTAATGGAAGAAAAAGAAATAGCTTGAAGA 1282
Db 1234 AGAAGTAAAGTGTAGTCTCTGAAAGTCCCTTAATGGAAGAAAGAGGCAACTTAAAGA 1293
Qy 1283 AGACCATGAAGAAACAAAGTGTCTGTGTGATATTGAAACAAAGCATCTGTCTGATCA 1342
Db 1294 AGACCATGAAGAAACAAAGTGTCTGTGAGACGTTGAGCATAGGAATCCTGTGTCTGA 1353
Qy 1343 GGTAGGCTGCACATGTCCTCCCTCCAGCTGTGGTGGAGGAGAGAACAGTCTCAATCAA 1402
Db 1354 GGTAGTGTGTGCACATGGGCACTCCGGGCTGTGTGGAGGAGAGACGCTGTCAATCAA 1413
Qy 1403 ACTTGAGATTTTGGAGGAAGCTCCAGAGAGAGAGGCTTCCAGCGTGGACTTGAAGA 1462
Db 1414 ACTTGTGTAGCTGGAGAGGCTCCGAGCGAGAGCGGCTTCCA---TGACCTGAAGA 1470
Qy 1463 GGAACACCATAGATAGCAACCGTGAATGGTGCAGTGCAGTGTCCCTAATGGGAACCTTGT 1522
Db 1471 GGAGACCATAGACAGACCAATCAATGTGTGAGTGTGCTTAAATGGGAACCTTGT 1530
Qy 1523 CCAGTTTCAAGCCGTGAGAACCAATTAACCTCCAGTGGCCACTACAGTATCACAC 1582
Db 1531 TCAGTTTCAGTCAAACTGTGAGCAACCATCAACTCCAGTGGCCACTATCAGTATCACAC 1590
Qy 1583 CGTGCAATAGGATTCGGGCTGTACAAAGAGCTACTCCATAAATTAATCTTGGCCAAAGT 1642
Db 1591 CGTGCAACAGGATTCGGCTGTGACAGGAGCTGTCCATTAATGATTAATCTGCGCAAGT 1650
Qy 1643 GGAGATTCGATGGGAGACTCCGGTGAACAAACCCCTTAAGCGCGCAATAATAGCTATACTTC 1702
Db 1651 GGGAGACTGATGGGAGATTCGGGCAACAGCCCTTGGAGACGCAACACAGCTTACACTTC 1710

Qy 1703 CTATACCATGCAATATGTGGCATGCCCTCTGGATTCAATTCGTTGCCAAAGAGGTGAACA 1762
Db 1711 CTACACTATGGCAATATGTGGCATGCCCTCTGGATTCAATTCGTTGCCAAAGAGGTGAACA 1770
Qy 1763 GAAGGGCGAAGAAATGGAGAGCTGACATGCGCTTAATCAGACTCCAAAGAGCGAATTCG 1822
Db 1771 AAAGGAGATGAATGGAGAGCGTGCATGCGCTAATGCAGATACCAAGAGCGAATTCG 1830
Qy 1823 AATGACAGTTTACACCACTTACTGCAATGCTGTGTCTGACCTTCACTCAGCATCTGAGAT 1882
Db 1831 AATGACAGTTTACACCACTTACTGCAATGCTGTGTCTGACCTTCACTCAGCATCTGAGAT 1890
Qy 1883 AGACATGAGTGTCAAGGCGAGAGATGGGTCTAGGTGACAGAAAGGAAGTAATGGCTCTCT 1942
Db 1891 GGACATGAGTGTGAAGGCTGAGATGGGCTGGGTGACAGAAAGAGAGCAGTGGCTCTCT 1950
Qy 1943 AGAAGATGGTATGACCAAGATAAGCTGAAAGTCTCTCTCTCTTCCAGATTCCTCGAGAT 2002
Db 1951 TGAAGATGGTATGACCAAGATAAGCCTGAAGTGTCCCTTCTCTTCCAGTTCCTCGAGAT 2010
Qy 2003 CCTTACAGCTGCTTTGGGTCAATTCGCCCATGGTGGCAATGACGTGAAGCAATGCCATTCG 2062
Db 2011 CCTTACAGCTGCTTTGGGTCAATTTGCCCATGGTGGCAATGACGTGAAGCAATGCCATTCG 2070
Qy 2063 GCCTCTGTGCTTTATATTTGGTTTATGACACAGAGATGTTCTTCAAAAAGTGGCAAC 2122
Db 2071 CCCTCTGTGCTTTGATCTTTGTTAT---AAACAAGAGCCTCTACAAAAGCGGCAAC 2127
Qy 2123 ACCAATATGGCTTCTACTTATGGTGTGTTGGTATCTGTGTGCTGTGGGTTTGGGG 2182
Db 2128 ACCCATATGGCTTCTGCTTTATGGTGTGTTGGCATTTGCAATGGGCTGTGGGTTTGGGG 2187
Qy 2183 AAGAAGAGTTATCAGACATGGGGAAGATCTGACCGATCAACCCCTCTAGTGGCTT 2242
Db 2188 AAGAAGAGTTATCAGACATGGGGAAGAGCCGTGACCCCAATCACACCTTCCAGTGGT 2247
Qy 2243 CAGTATTAACCTGGCATCTGCCCTCACCTGTGTGTGATGTCATCAAAATATTGGCTTCCCAT 2302
Db 2248 CAGTATTAACCTGGCGTCTGCTTAACTGTGTGTGATCGATCAAACTATTGGCTTCCCAT 2307
Qy 2303 CAGTACAAACATTTGAAAGTGGCTCTGTGTGTGCTGTGTGGCTGGCTCCGCTCAAGAA 2362
Db 2308 CAGCACAAACATTTGAAAGTGGCTCTGTGTGTGCTGTGTGGCTGGCTCCGATCAAGAA 2367
Qy 2363 GGCTGTGACTGGCGTCTCTTTGTAACATTTTATGGCTGGTGTGTCACAGTCCCCAT 2422
Db 2368 GGCTGTGACTGGCGACTGTTTTCGAAACATTTTATGGCTGGTGTGTCACGGTCCCCAT 2427
Qy 2423 TTCTGGAGTTATCAGTGTGCTGCATATGGCAATCTTCAAGATATGTCATCTCAGAAATGTG 2482
Db 2428 CTCTGGGTTATCAGTGTGCGCTATCATGGCAGTATTCAAGTACATCATCTCCAGTGTG 2487
Qy 2483 AAGCTGTTTGAATTTAAATTTGTCTCAATGTTTGGGACCATCTTTAGGTATTTCTGCTCC 2542
Db 2488 ACGCT---GGGGTTGAAAGCTGTCTAGTGTCTGGACCATTTGTACACATTTCTG-TTC 2542
Qy 2543 CCTGAAGAAATGATTAAGTGTAAACAGAACTGACAGAGTCTTTTATTTTGGGAGCCA 2602
Db 2543 CTAGAGAAACGCTCAGAGTGT-TGCTGAAGACAGGCAAGGGTCT----TAAAGGAGCGG 2596
Qy 2603 GAGGAGGGAAGTGTACTTGTCTATTAAGTGTGCTTTTGTGCTAAATATGAAATTTGTCTCAA 2662
Db 2597 TGGGAAGGAAGTAAATTTACACTAATATGCTTTTGTGCTAAATATGACTTATCTCAA 2656
Qy 2663 ATTAGCTGTGAAAAATAGCCGGGTTCACATGGCTCTGCTGAGGTGCCCTTTCTCTTCG 2722
Db 2657 ATTAGCTGTAAAAATAGCAGGTTTCCATTTCAATTCATTTCCAAAGT-CCCTTTTCTCTG 2715
Qy 2723 GGCTGTGAATTCCTGTACATATTTCTACTTTTGTATCAGGCTTCAATTTCCATTTCT 2782
Db 2716 GGCTATGAATTCCTGTACATATTTCTCTAC-TTTTGTATCAGGCTCAATTCAGTATGT 2774


```
QY 2783 TTTAATGTTGCTCTGCAAGTACTGTGATTTTTTTTTTTTCTTTTTTTTTTTTAAACATGAGA 2842
Db 2775 TTTAATGTTGCTCTGCA-----GATACTTAGTGGTCTCTTTTAAACAGCCAGCA 2826
QY 2843 GCGGTTTGACAGAGTCTCTGCGTTGTTGTTTCCACAGCTTCTGCGCTCACATGCA 2902
Db 2827 GAGCCATTGATGCGCATGCTGCTTTTGGCTCACCAGCTTCTTCCCAACATGCA 2886
QY 2903 AGGATTAAACAAACAAATAATAACATCAACCTTCCCTTGTAGTCTCTTATATAAGTAGAG 2962
Db 2887 AGGATTAAACAAACATGTAAGCTTCCCTCCCTCATAGTCTCTCATAGAAATAGTC 2946
QY 2963 TCCTTGGTACTCTGCGCTCTCTGCTAGTGGAGATCTATTTGGCATATTCGGAGCTT 3022
Db 2947 AC---GGCACTCTGCTC-CCTGTCACATGAGGAGGTTCTGTGG---ATGTGTGACAACTT 3000
QY 3023 CTTAGAGGATGAGGTTCTTTGAACACAGTGAATAATTTAAAT---TAGTAACTTTTTTTCG 3079
Db 3001 CTTAGAGGCGGAGATCTTTG-GCAGTGGNAATATAAGTTTGTAGTAACCTTTTTCG 3059
QY 3080 AAGCAGTTTATTGACTGTTATTGCTAAGAGAGTGAAGAAAGAAAGCCCTGTGGCAAT 3139
Db 3060 AAACAGTTTACAGCAC--ATGTTGCTAAGAGCAGGAGACAAA-----GCCCTGCGGT 3112
QY 3140 CTTGTTATTCTTTAAGATTTTTCGACGTGTTGGATGATGATGAAGTGAATGTGA 3199
Db 3113 TGTGGTATTCTTCTGAGATTTTTCGACGTGTTGGATGATGATGAAGTGAATGTGA 3172
QY 3200 CTTTGGGCAAGTTAAATGGACAGCTTCCATGTTTCATTTGTTACCTCTTAACCTGAATA 3259
Db 3173 CTTTGGGCAANTCAATGGACAGCTTCCATGTTTCATGTTTCATCTTACCTCTTAACCTGAAT- 3231
QY 3260 AAAAAAGCCTACAGTTTTTGAAGAAAAAAA 3288
Db 3232 AAAAAAGCCTACAGTTTTTGAAGAAAAAAA 3260

RESULT 12
AB000489
LOCUS Rattus norvegicus mRNA for RPHO-1, complete cds.
DEFINITION Rattus norvegicus adult male parathyroid glands tissue lib:lambda
ACCESSION AB000489
VERSION AB000489.1 GI:9695267
KEYWORDS Rattus norvegicus adult male parathyroid glands tissue lib:lambda
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (sites)
AUTHORS Tatsumi,S., Segawa,H., Morita,K., Haga,H., Kouda,T., Yamamoto,H.,
Inoue,Y., Ni,T., Kawai,K., Taketani,Y., Miyamoto,K. and Takeda,S.
TITLE Molecular cloning and hormonal regulation of Pit-1, a
sodium-dependent phosphate cotransporter from rat parathyroid
glands
JOURNAL Endocrinology 139 (4), 1692-1699 (1998)
MEDLINE 98187800
REFERENCE 2 (bases 1 to 2890)
AUTHORS Miyamoto,K.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1997) Ken-ichi Miyamoto, University of Tokushima,
School of Medicine, Department of Clinical Nutrition; 3-18-15
Kuramoto-cho, Tokushima, Tokushima 770-8503, Japan
(E-mail:miyamoto@nutr.med.tokushima-u.ac.jp, Tel:81-886-33-7095,
Fax:81-886-33-7094)
FEATURES
source
1..2890
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="parathyroid glands"
/dev_stage="adult"
```

CDS

```
/tissue_lib="lambda gt10"
92..2137
/codon_start=1
/protein_id="RPHO-1"
/db_xref="GI:9695268"
/translation="MASTLAPITSTLAAVTASAPPKYDNLMLIIGLIIAFVLAFLSVG
ANDVANSFGTAVGSGVTLKQACILASIFETVGSALLGAKVSETIRGLIDVKNYAT
QDLIMAGSVAMFGSAGVQLVFLKPISGTHCTVIGATIGSLVAKGQEGKWSBLI
KIVMSWFLVPLSGIMSGILPLFRALRKPVNGLRALPIFAVACTGIGNLPSIM
YTGAELGFDKLPKAGTILISGCAVFCALIVWFVCPRMKEIREVVKSSSESPLM
EKKXNLKHEETKMAPDVENRNPNVSEVVCATPLRAVVEERTVSKFLGDLSEAPRE
LLPMLKBEETNIDGTINGAVLPNGNLVQFQTSVNSQINSSGHTYQTHVHKDSGLYKE
LLHKHLAKRGVDCMGSDGKPLRRNNSYTSYMAICGMLPDSFRAKEGQKQDEWDL
TWPNADTKRIRMSDYSYCNVADSHSEMDMSVKAEMGLGDRKSGSGSELEWYDQ
DKPEVSLPQFLQILATCFGSAHGNVDVNAIGPLVALYLVTRETDVTTKATPIWL
LLYGVGICMLGWVGRVIOITMGKDLPTIPSSGFSIELASAFVTVVNASNIGLPIST
THCKVGSVSVGVGLSKKAVDWRLFRNFMFVTVPISGVISAAIMAVFKHILPV"
BASE COUNT 710 a 631 c 722 g 827 t
ORIGIN
Query Match 59.4%; Score 1954; DB 10; Length 2890;
Best Local Similarity 84.7%; Pred No. 0;
Matches 2445; Conservative 0; Mismatches 380; Indels 63; Gaps 20;
QY 395 TTGAAGGCGCTCAGTAGTTCCTTTA----CTAAACAACCACTACTTCCAGAGAATGGCA- 449
Db 39 TGAAGGCGCTCAGTAATTCGTATTATTTTAAAGAACCACTACACTACAGATGGCAT 98
QY 450 -----ACGGTGATTACAGTACTACAGCTGCTACCGCCGCTCTCTGTCTCTTGGTG- 500
Db 99 CTACCCCTGGCCCCGATTACTAGTACTAGCTGCTTACTGCTTCTTCTCCACCGAAGT 158
QY 501 --GACTACTATGATGCTCATCCGGGCTTCATTTATTTGCTATTTGCTTTGGCATTCCTCG 558
Db 159 ATGACAAATCTGGGATGCTCATCTCGGGCTTCATTTGCTATTTGCTTTGGCATTCCTCG 218
QY 559 TGGGAGCCAAATGATAGCAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCC 618
Db 219 TGGGAGCCAAATGATAGCAAAATCTTTTGGTACAGCTGTAGGCTCAGGTGTAGTGACCC 278
QY 619 TGAAGAGCCTGCATCTCTAGTACGATCTTTGAAACAGTGGGCTCTGTCTTCTACGGGG 678
Db 279 TGAAGAGCCTGCATCTCTAGTACGATCTTTGAAACCTGTGGGCTCCGCTTGTGGGG 338
QY 679 CCAAGTGAGCGAAGAACCATCCGGAAGGCTTCATTTGAGTGAGATGTACAACTGCATC 738
Db 339 CCAAGTGAGTGAAACCATCCGGAAGGCTTCATAGATGTGGAGAGTACAACTGCATC 398
QY 739 AAGGGCTACTGATGGCGGCTCAGTCAGTCTATGTTTGGTTCCTGTGTGGCAACTCG 798
Db 399 AAGATCTGCTCATGCTGGCTGCTGTCAGTCTATGTTTGGTTCCTGTGTGGCAACTAG 458
QY 799 TGGCTTCTGTTTGAAGCTCCCTATTTCTGGAACCATTTGATTTGTTGGTGAACATTTG 858
Db 459 TGGCTTCTGTTTGAAGCTTCTCTATTTCTGGACCCATTTGATTTGCTGGTGAACCATTTG 518
QY 859 GTTCTCTCCCTCGTGGCAAGGGGAGGAGGTGTCAGTGGTCTGAACTGATAAAAAATG 918
Db 519 GTTCTCTCCCTTGTGGCAAGGGAGCAAGAGGGGTATCAATGGTCCGAACATGATAAAAAATG 578
QY 919 TGATGCTTGTGTTTGGTCTCCCTGCTTTCTGGAAATATGCTGGAATTTTATTTCTTCC 978
Db 579 TGATGCTTGTGTTTGGTCTCCCTGCTTTCTGGTATATGCTGGAATTTTATTTCTTCC 638
QY 979 TGGTTCGTGATTCATCTCCATAAGGAGATCCAGTTCCTTAATGTTTGGAGCTTTGCG 1038
Db 639 TTGTTTCGTGATTCATCTCCGTAAGGAGATCCAGTTCCTTAATGCTTACGAGCTTTAC 698
QY 1039 CAGTTTTCTATGCTTCACAGTTGGAATAAACCTCTTTTCCATCATGTATCTAGTGAGCAC 1098
Db 699 CCAATTTTATGCTTCGCAATTTGGAATCAACCTCTTTTCCATTATGTTATCTAGTGAGCAC 758
```


Db 2875 AATAAAAA 2882

RESULT 13
CGUI13946

LOCUS CGUI13946 2040 bp mRNA linear ROD 09-SEP-1994

DEFINITION Cricetus griseus gibbon ape leukemia retrovirus receptor/sodium-dependent Pi transporter mRNA, complete cds.

ACCESSION U13946

VERSION U13946.1

KEYWORDS Gi:535457

SOURCE Cricetus griseus.

ORGANISM Cricetus griseus

REFERENCE 1 (bases 1 to 2040)
Wilson, C.A., Farrell, K.B. and Eiden, M.V.
Properties of a unique form of the murine amphotropic leukemia virus receptor expressed on hamster cells
J. Virol. 68 (12), 7697-7703 (1994)

AUTHORS Eiden, M.V.

TITLE Direct Submission

JOURNAL Submitted (24-AUG-1994) Maribeth V. Eiden, Laboratory of Cell Biology, National Institute of Mental Health, Bldg. 36 Convent Dr. MSC 4068, Bethesda, MD 20892-4068, USA

FEATURES
source location/Qualifiers
1..2040
/organism="Cricetus griseus"
/db_xref="taxon:10029"
/clone="pTA-BGR"
/sex="male"
/cell_line="E36"
/cell_type="fibroblast"
/tissue_type="lung"
/germline

CDS
1..2040
/codon_start=1
/product="gibbon ape leukemia retrovirus receptor/sodium-dependent Pi transporter"
/protein_id="AAA57033.1"
/db_xref="GI:535458"
translations="MATLITTAATAASGPLVDYLMILGLIFAFVLAFLSVGVANDVANSFGTAVGSGVTLTKQACILASIFETVGSVLTKGAKVSETIRKGLIDVEMYNSTQGLSWFVSLSSGIMSGLIFLVRAPIHKADVPVNGLRALPVFACTVGINLFSIMYTKA PLIGFDKLPWGLTILSVGCAVICALIWMFVFCPRMKRTIERIKGSPLESIMKKN NLKDEHETKMGIDVNGSVSEVACATGLOAVVEERTVSPKLGDLSEAPERERLP SLDLKETSIDSTINGAVOLPNGNLVQVSTVSNQINSSGHYQYHTVHKDSGLYKELL HKLHLKRGVGDGSGYKPLRRNNTSYTMAICMPLDSFRAGEQGEENEKULTV PNADEKRRIMHDSYTSYCNVAVSDLSHASEIDMSVKAEMGLGDRKSGNSLEGWYDQDLPEVSLVQFQILITACFGSPAHGNDVSNAGLPVALYLVYDTGDSVSKVATPIWLL YVGWGCIVGLWVGRVVIOTMGKDLPTPTSSGFSIELASALTWVIASNIGLPISITTH CKVGSVSVGWLRSKKAVDWRDLFRNIFMAFWFVTPISGVISAAMAIFRYVILRM"

BASE COUNT 492 a 446 c 529 g 573 t

ORIGIN

Query Match 58.8%; Score 1934.4; DB 10; Length 2040;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1974; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 444 ATGCAACGGTGATTACAGTACTACAGCTGCTACCGCCGCTTCTGGTCTTTGGTGGAC 503
|||||

Db 1 ATGCAACGGTGATTACAGTACTACAGCTGCTACCGCCGCTTCTGGTCTTTGGTGGAC 60
|||||

QY 504 TACCTATGGATGCTCATCTCGGGCTTCATTATTGCAATTCGTTCGTGGCATTCCTCGTGGGA 563
|||||

Db 61 TACCTATGGATGCTCATCTCGGGCTTCATTATTGCAATTCGTTCGTGGCATTCCTCGTGGGA 120
|||||

QY 564 GCCAATGATGTAGCAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCCCTGAAG 623
|||||

538	TTGGTTCAATCTCTCCACGTGCTTTCTGGTATTAATGATCTGGAAATTTATTAATCTTCCTCTGTGTGC	597	
Qy	986	TGCATTTCAATCTCCATAAAGCAGATCCAGTTTCTTAATGGTTTGGAGCTTTTGCCAGTTTT	1045
Db	598	TGCGTTCAATCTCCGTAAGACAGATCCGGTTCTTAATGGTTTGAGAGCTTTGGCCAGTTTT	657
Qy	1046	CTATGCGCTGCACAGTTGGAAATAAACCTCTTTTCCATCATGTATATCTGGAGCACCTTTGCT	1105
Db	658	CTATGCGCTGCACAGTTGGAAATAAACCTCTTTTCCATCATGTATATCTGGAGCACCTTTGCT	717
Qy	1106	GGGCTTTTGACAAACTTCTCTGTGGGGTACCATCCTCATCTCGGTGGGATGTGCAGTTTT	1165
Db	718	GGGCTTTTGACAAACTTCTCTGTGGGGCACCATCTCATCTCGGTGGGATGTGCAGTTTT	777
Qy	1166	CTGTGCCCTTATCGTCTGTGGTTCTTTTGATGTCCAGGATGAAGAGAAAAATTTGAACGAGA	1225
Db	778	CTGTGCCCTTATCGTCTGTGGTTCTTTTGATGTCCAGGATGAAGAGAAAAATTTGAACGAGA	837
Qy	1226	AATAAAGTGTAGTCTCTTCTGAAAGCCCTTAAATGGAATAAAGAAATAGCTTTGAAAGAAGA	1285
Db	838	AATAAAGTGTAGTCTCTTCTGAAAGCCCTTAAATGGAATAAAGAAATAGCTTTGAAAGAAGA	897
Qy	1286	CCATGAAGAAACAAAGTTGCTGTGTGGTGATATTGAAACAAGCATCCTGTCTTGAGGT	1345
Db	898	CCATGAAGAAACAAAGTTGCTGTGTGGTGATATTGAAACAAGCATCCTGTCTTGAGGT	957
Qy	1346	AGGCGCTGCCACTGTGCGCCCTCCAGGCTGTGGTGGAGAGAGACAGTCTCATTTCAAAC	1405
Db	958	AGGATCTGCCACTGTGCGCCCTCCGGGCTGTGGTGGAGAGAGACAGTCTCATTTCAAAC	1017
Qy	1406	TGGAGATTTTGGAGGAAGCTCCAGAGAGAGAGAGCGTTCCAGCGTGGACTTTGAAAGAGGA	1465
Db	1018	TGGAGACTTTGGAGGAAGCTCCAGAGCGAGAGAGCGTTCCAGTGTGGACCTGAAAGAGGA	1077
Qy	1466	AACCAGCATAGATAGCACCGTGAATGGTGCAGTGCAGTTGCCTTAATGGGAACCTTTGTCCA	1525
Db	1078	AACCAGCATAGACAGCGCCATGAATGAGAGCAATTCAGTTGCCTTAATGGGAACCTTTGTCCA	1137
Qy	1526	GTTCAGTCAAGCGCTCAGCAACCAATAAATCCAGTGGCCACTACCAGTATCACACCGT	1585
Db	1138	GTTCATATCAGCGCTCAGTAACCAAGATGAATCCAGTGGCCACTATCAGTATCATACCGT	1197
Qy	1586	GCATAAGGATTTCCGGCTGTACAAGAGCTACTCCATAAATTAATCTTGCCAAAGGTGGG	1645
Db	1198	GCATAAAGATTTCTGGCTTATACAAGAGCTGCTGCATAAATTAATCTTGCCAAAGGTGGG	1257
Qy	1646	AGATTGCATGGGAGACTCCGGTGACAAACCCCTTAAGGGCAATAATATAGCTATACTTCCCTA	1705
Db	1258	AGACTGCATGGGAGACTCTGRCACAAACCCCTTGAGGCGCAATAATATAGCTATACTTCCCTA	1317
Qy	1706	TACCATGGCAATATGTGGCATGCTCTGGATTCAATCCGTGCCAAAGAAAGGTGAACAGAA	1765
Db	1318	CACCATGGCAATTTGTGGCATGCTCTGGATTCAATCCGTGCCAAAGAAAGGTGAACAGAA	1377
Qy	1766	GGGCGAAGAAATGGAGAGCTGCATAGCCCTTAATGACAGACTCCAAGAAAGCGAATTCGAAT	1825
Db	1378	AGGTGAAGAAATGGAGAGCTGCATAGCCCTTAACCGCGAGAGCAAGACGAATTCGAAT	1437
Qy	1826	GGACAGTTTACACAGATTACTGCAATGTGTGTCTGAACCTTCACTCAGCATCTGAGATPAGA	1885
Db	1438	GGATAGTTTACACCAAGTTATTGCAATGTGTGTCTGTGACATTCACCTCGGCATCCGAGATGGA	1497
Qy	1886	CATCAGTGTCAAGSCAGAGATGGGCTTAGGTGACAGAAAGGAGTAATGGCTCTCTAGA	1945
Db	1498	CATCAGTGTCAAGSCAGAGATGGGCTTAGGTGACAGAAAGGAGTAATGGCTCTCTAGA	1557
Qy	1946	AGAAATGGTATGACACAGGATAGCCCTGAAGTCTCTCTCTCTTCCAGTTCTCTGCAGATCCCT	2005
Db	1558	AGAAATGGTATGACACAGGATAGCCCTGAAGTCTCTCTCTCTTCCAGTTCTCTGCAGATCCCT	1617
Qy	2006	TACAGCTGCTTTGGGTCATTTGCGCCATTTGGCCAATGACGTAAAGCAATGCCATTTGGGCC	2065
Db	1618	GACAGCTGCTTTGGGTCATTTGCGCCATTTGGGAAATGATGTACAGAAATGCCATTCGGTCC	1677

Qy	2066	TCTGGTTCCTTTATATTTGGTTTATAGACACAGAGATGTTTCTTCAAAGTGGCAACCC	2125
Db	1678	TCTGGTTCCTTTATATCTGCTTTTATGACACAGAGATGTTTCTTCAAAGTGGCAACCC	1737
Qy	2126	AAATATGGCTTCTACTCTATGGTGGTGTGGTATCTGTGTTGGTCTGTGGGTGTGGGGAG	2185
Db	1738	GATATGGCTTCTGCTCTATGGTGGAGTTGGTATTTGTATTTGGTCTGTGGGTGTGGGGAAG	1797
Qy	2186	AAGAGTTATTCAGACCAATGGGGAAGGATCTGACACCGATCACACCTCTATAGTGGCTTCAG	2245
Db	1798	GAGAGTTATTCAGACCAATGGGGAAGGATCTGACACCGATCACACCTTCTAGTGGCTTCAG	1857
Qy	2246	TATTGAACCTGGCATCTGCCCTCACTGTGGTGATTGCATCAAAATATTGGCCTTCCCATCAG	2305
Db	1858	TATTGAACCTGGCATCTGCCCTCACTGTGGTAAATTGCATCAAAATATTGGTCTTCCCATCAG	1917
Qy	2306	TACAACACATTGTAAGTGGGCTCTGTGTGTCTGTGTTGGCTGGCTCCGGTCCAAGAAGGC	2365
Db	1918	TACAACGCAATTGTAAGTGGGTTCTGTGTGTATCTGTTGGCTGGCTCCGATCCAAAAGGC	1977
Qy	2366	TGTTGACTGGGTCCTCTTTCGTAACATTTTATATGGCTGGCTGGTTGTGCACAGTCCCATTTTC	2425
Db	1978	TGTTGATTGGCGACTCTTTCGCGAACATTTTATATGGCTGGTGTGTTACAGTCCCATTTTC	2037
Qy	2426	TGGAGTTATCAGTGTGCGCCATCATGGCAATCTTTCAGATATGTCATCTCTCAGAATGTGAAG	2485
Db	2038	TGGCGTTATCAGTGTGCTGCTATTATGCGACTCTTCAAATATGTCATCTCTGACGGTGAAG	2097
Qy	2486	CTGTTTGAGATTAATAATTTGTGTCAATGTTT	2516
Db	2098	CTGTTTGAGATTAATAATTTCTGTCAATGTTT	2128

RESULT 15
AF063024
LOCUS AF063024 2503 bp mRNA linear ROD 05-MAY-1999
DEFINITION *Cricetulus griseus* phosphate transporter 1 mRNA, complete cds.

KEYWORDS
SOURCE
ORGANISM
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetus.
REFERENCE
1 (bases 1 to 2503)
AUTHORS
Chaudry, G.J., Schmitz, C., Ting, Y.-T., Farrell, K.B., Lie, Y.S.,
Petrooulos, C.J. and Eiden, M.V.
TITLE
Isotypes of the Chinese hamster receptors for gibbon ape leukemia
virus have dramatically different infection permissivities
JOURNAL
Unpublished
2 (bases 1 to 2503)
AUTHORS
Petrooulos, C.J., Lie, Y.S. and Schmitz, C.
TITLE
Direct Submission
JOURNAL
Submitted (04-MAY-1998) Cell Banking and Characterization,
Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
FEATURES
Location/Qualifiers
source
1. .2503

```

CDS
231..2273
      cdsl_name= CHO-K1
      /notes="Gibbon ape leukemia virus receptor"
      /codon_start=1
      /product="phosphate transporter 1"
      /protein_id="AAD28692.1"
      /db_xref="GI:4741732"
      /translation="MASTVAPITNTLTATVAAAASVDDKVMVILGFIIFAVLAFSVG
      ANDVANSFGTAVSGVWVTLKQACILASIFETVSGALLGAKVSETIRKGLIDVMYNET
      QDLLMAGSVASGMSAVWLKPSFKLIPSGTHICVATGIFSLVAKGGQGVKWSGLI
      KIVMSFVSPGLNSGIFLFLVRLIRUKADPVGNGRLAPVFVACTIGNLNSLIM
      YTCAPILGDKPLPLWCTLTILSVGCAVFCALFYVWFPGCEPMKRKTRTEIKSSSESPLM

```

EKKNNLKEDHEETKMSLGDVNGNSVSEVACATGPLRAVVBERTVSPFKLGDLEBAPER
ERLPLDLKEEESIDGTINGAVOLPGLNVLQPSQTVSNQINSNGHVQYVTHKDSGLXK
ELUHLKLAUKVDCMGSDGDKPLRRNNSYTSYMAICGMLPDSFRKSGEQKESBET
LTHPVDYTKRIRMDSDYSYCNVADLHSESEMNMVNAEMGNRRGSLSEEWCDQD
KPEVSLQFQLOILTACFGSAHGNNDVSNATGLPVLYLVVYTDQVSTKATPTPWL
LYGGVIGICIGLWVWRRVIOQMGKOLPTIPSSGFSIELASALTUVVANSIGLPISTT
HCKVGSVSVGVHLRSKKAVDMRLPFIWVPTVPIGVSIAAIMAVKFIILKV"

BASE COUNT 624 a 549 c 615 g 715 t
ORIGIN

Query Match 55.3%; Score 1820.2; DB 10; Length 2503;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 2096; Conservative 0; Mismatches 253; Indels 34; Gaps 9;

QY	215	TCGTCTCGTGCTTTTAGCCCTCTGAGCCAAAGAAACCCAGACACAGATGCCCATAC	274
DB	1	TCGTCTCGTGCTTTTAGCCCTCCAGAGCCAAAGAAACCCAGACACAGATGCCCATAC	60
QY	275	GCAGCGTATAGCAGTAACCTCCAGCTCGGTTCTGTGCGGTAGTTTACAGTATTAAAT	334
DB	61	GCAGCGTATAGCAGTAACCTCCAGCTCGGTTCTGTGCGGTAGTTTACAGTATTAAAT	120
QY	335	TTATATAATATATATTTATATAGCATTTTGTATACCTCATATTTCTGTTTACACATC	394
DB	121	TTATATAATATATATTTATATAGCATTTTGTATACCTCATATTTCTGTTTACACATC	177
QY	395	TTGAAGGGCGCTAGTAGTTCTCTT-----ACTAAACACACTACTCCAGAGAAATGGCA	449
DB	178	TTGAAGGGCGCGAGTAATCTCTTATTAAATTAAGAACCACTACACTAAAAAATGGCAT	237
QY	450	-----ACGGTGATTACAGTACTACAGCTGCTACCGCGCTCTGCTCTTGGTGG	501
DB	238	CTACAGTGGCAGCATTAATACTCTAGCTACTGTTGCTACTGCTGCTGCTGCTGCTG	297
QY	502	ACTAC---CTATGGATGCTCATCTCGGCTTCATTATGTCATTTGCTTGGCATCTCGG	558
DB	298	ATGACAAAGTATGGATGGTCATCTCGGCTTTATCATTCGATTTGCTTGGCATCTCGG	357
QY	559	TGGAGCCAAATGATGTAGCAAAATCTTTTGGTACAGCTGTGGCTCAGGTGTAGTACCC	618
DB	358	TGGAGCCAAATGATGTAGCAAAATCTTTTGGTACAGCTGTGGCTCAGGTGTAGTACCT	417
QY	619	TGAAGCAAGCTGCTATCTAGCTAGCATCTTTGAAACAGTGGCTCTGCTTACTTGGGG	678
DB	418	TGAAGCAAGCTGCTATCTAGCTAGCATCTTTGAAACAGTGGCTCTGCTTACTTGGGG	477
QY	679	CCAAAGTGACGAAACCATCCGAGGGCTTGATTGACGTGGAGATGTACAACTCGACTC	738
DB	478	CCAAAGTGAGTGAACCATCCGAGGGCTTGATTGATGTGGAGATGTATAACGAAACTC	537
QY	739	AAGGGCTACTGATGGCGGCTCAGTCACTGCTATGTTGGTTCGTGCTGTGGCAACTCG	798
DB	538	AAGACTTGCTATGGCTGGCTCGCTCAGTCTATGTTGGTTCGTGCTGTGGCAACTAA	597
QY	799	TGCTTCTGTTTTTGAAGCTCCCTATTCTTGGAAACCATTTGATTGTTGGTGAACCTATTG	858
DB	598	TGCTTCTGTTTTTGAAGCTGCCCATTTCTGGHACCATTTGATTGTTGGTGCACCATTTG	657
QY	859	GTTTTCTCCCTCGTGGCAAGGGGCGAGGGGTGCAAGTGGTCTGAACTGATAAAATTTG	918
DB	658	GTTTTCTCCCTCGTGGCAAGGGGCAAGGGTGTCAAGTGGTCTGAGCTGATAAAATCG	717
QY	919	TGATGTCTTGGTTCGTGCTCCCATCTGTTTCTGGAAATTTATGCTGGAATTTTATCTTCC	978
DB	718	TGATGTCTTGGTTCGTCTCTCCACTGTTTCTGGTATTATGCTGGAATTTTATCTTCC	777
QY	979	TGCTTCTGCTATCTCCATAAGCAGATCCAGTCTCTATGTTTGGAGCTTTTCG	1038
DB	778	TTGTTCTGCTATCTCTCCGTAAGCAGATCCAGTCTCTATGCTTCTGCTGCTGCTTAC	837
QY	1039	CAGTTTTCTATGCTGCACAGTTGGAAATAACCTCTTTTCCCATCTATGCTATGAGCAC	1098
DB	838	CGGTTTTTATGCTGTACAAATTTGGAATAAACCTCTTTTCCATTTATGATATCTGGAGC	897

QY	1099	CGTTCTCGGCTTTGACAAATTTCTCTGTGGGGTACCATTCTCATCTCGTGGGATGTG	1158
DB	898	CGTTCTCGGCTTTGACAAATTTCTCTGTGGGGTACCATTCTCATCTCGTGGGATGTG	957
QY	1159	CAGTTTTCTGTGCCCCCTTATCGTCTGTGGTTCTTTGTATGTCCCAGGATGAAGAGAAAATTTG	1218
DB	958	CAGTTTTCTGTGCCCCCTTATCGTCTGTGGTTCTTTGTATGTCCCAGGATGAAGAGAAAATTTG	1017
QY	1219	AACGAGAAATTAAGTGTAGTCTCTTCTGAAAGCCCTTAATGGAAAAAAGAAATAGCTTGA	1278
DB	1018	AACGAGAAATTAAGTGTAGTCTCTTCTGAAAGCCCTTAATGGAAAAAAGAAATAGCTTGA	1077
QY	1279	AAGAGACCATCAAGAAACAAAGTTGCTGTGGTGATATTGAAACAAAGCATCTGTTT	1338
DB	1078	AAGAGACCATCAAGAAACAAAGTTGCTGTGGTGATATTGAAACAAAGCATCTGTTT	1137
QY	1339	CTGAGTGTAGGCGCTGCCACTGTGCCCCCTCCAGGCTGTGGTGAGAGAGAAACAGTCTCAT	1398
DB	1138	CTGAGTGTAGTGTGCCACTGTGGCCCTCTCCGGCTGTGGTGAGAGAGAAACAGTCTCAT	1197
QY	1399	TCAAACTTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTTGA	1458
DB	1198	TCAAACTTTGGGAGCCCTGGAGGAAGCTCCAGAGCGAGAGCGGCTTCCCA---TGGACTTGA	1254
QY	1459	AAGAGAAACAGCATAGTACGACCGTGAATGTTGTCAGTGCAGTTGCTTAATGGGAACC	1518
DB	1255	AGGAGGAGCAGCATAGTATGTTACCAATGTTGGTGGAGTGGCTTAATGGGAACC	1314
QY	1519	TTGTCAGTTCAGTCAAGCCGTCAGCAACCAAAATAAATCCAGTGGCCCACTACCAATATC	1578
DB	1315	TTGTTTCACTTCACTCAAACTGTACCAACCAAGTCACTCCAGTGGCCCACTATCATATC	1374
QY	1579	ACACCGTGCATAAGATTTCCGGCTGTACAAAGAGCTACTCCATTAATATCATCTTGGCA	1638
DB	1375	ACACTGTGCACAAAGATTTCTGGCTTTGTACAAAGAGTTGCTTTCATTAATATCATCTTGGCA	1434
QY	1639	AGTGGGAGATTTGTCATGGGAGACTCCGGTGACAAACCTTAAGGCGCAATATAGCTATA	1698
DB	1435	AGTGGGAGATTTGTCATGGGAGACTCTGGGATTAAGCCGTTGAGGCGCAACACAGCTACA	1494
QY	1699	CTTCTCTATACCATGGCAATATGTGCATGCTCTGAGTTCATTTCCGTGCCAAAGAGGTG	1758
DB	1495	CTTCTCTACCATGGCAATATGTGCATGCTCTGAGTTCATTTCCGTGCCAAAGAGGTG	1554
QY	1759	AACAGAGGGCGAAGAAATGGAGAGCTGACATGCGCTTAATGCAGACTCCAGAGAGCGAA	1818
DB	1555	AACAGAGGGGAGAGAAACCGAGAGCGCTGACATGGCCTTAATGTAGATACTAAGAGAGCGGA	1614
QY	1819	TTTGAATGGACATTTACACCACTTACTGCAATGCTGTGCTGACCTTCACTCAGCATCTG	1878
DB	1615	TTTGAATGGACATTTACACCACTTACTGCAATGCTGTGCTGACCTTCACTCAGCATCTG	1674
QY	1879	AGATAGACATGAGTGTCAAGCGAGAGATGGGTCTAGGTGACAGAAAAGGAAGTAATGGCT	1938
DB	1675	AGATAGACATGAGTGTCAAGCGTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	1728
QY	1939	CTCTAGAGAAATGGTATGACCAAGGATAGCTGAAAGTCTCTCTCTCTCTCTCTCTCTCTC	1998
DB	1729	CTCTTGAAGAGTGGTGTGACCAAGGATAAGCCCTGAAAGTCTCTCTCTCTCTCTCTCTC	1788
QY	1999	AGATCTTTACAGCTGCTTTGGTGCATTTGCGCCCATGGTGGCAATGACAGTAAGCAATGCCA	2058
DB	1789	AGATCTTTACAGCTGCTTTGGTGCATTTGCGCCCATGGTGGCAATGACAGTAAGCAATGCCA	1848
QY	2059	TTGGGCGCTCTGCTTCTTATATTTGGTTTATGACACAGGAGATGTTTCTTCAAAAGTGG	2118
DB	1849	TTGGTCTCTGCTTCTTATCTTGTATCTTTTATACAACTCAAGATGTGTCTACAAAGATAG	1908
QY	2119	CAACACCAATATGGCTTCTACTATGTTGGTGTGTTGTTATCTGTTGTTGTTGTTGTTGTT	2178
DB	1909	CAACACCAATATGGCTTCTGCTTTTATGGTGGTGTGGCATCTGCAATAGGCTGTGGGTTT	1968

Qy 2179 GGGGAAGAAGATTATCCAGACCATTGGGGAAGGATCTCAGACCCGATCACACCCCTCTAGTG 2238
 Db |||||
 Qy 1969 GGGGAAGAAGATTATCCAGACCATTGGGGAAGGATCTCAGACCCGATCACACCCCTCCAGTG 2028
 Db |||||
 Qy 2239 GCTTCAGTATTGAAGTGGGCTGCGCTCACTGTGGTGAATGCGATCAAAATATTGGCCTTC 2298
 Db |||||
 Qy 2029 GTTTCAGTATTGAAGTGGGCTGCGCTCACTGTGGTGAATGCGATCAAAATATTGGCCTTC 2088
 Db |||||
 Qy 2299 CCATCAGTACACACATTGTAAAGTGGGCTCTGTGTGTCTGTGGCTGGCTCCGGTCCA 2358
 Db |||||
 Qy 2089 CCATCAGCACAACACACTGTAAAGTGGGCTCTGTGTGTCTGTGGCTGGCTCCGATCCA 2148
 Db |||||
 Qy 2359 AGAAGGCTGTGACTGGGCTCTCTTTCTGTAAACATTTTATGGCCTGGTTGTCAAGTCC 2418
 Db |||||
 Qy 2149 AGAAGGCTGTGACTGGGCTCTCTTTCTGTAAACATTTTATGGCCTGGTTGTCAAGTCC 2208
 Db |||||
 Qy 2419 CCATTTCTGGGATTATCAGTGTGCTGCCATCGCAATCTTCAGATATGTCATCCTCAGAA 2478
 Db |||||
 Qy 2209 CCATTTCTGGGATTATAGTGTCTCTATCATGGCAGTATTCAAATTCATCCTCGAAAG 2269
 Db |||||
 Qy 2479 TGTGAAGCTGTTTGAGATTAAATTTGTGTCAATTTTGGGACCATCTTAGGTATTCCTG 2538
 Db |||||
 Qy 2269 TGTGAAGC----TGGGGTTGAATTTTGTCAAGCATCTGGGACCATCACACATTCCTG 2324
 Db |||||
 Qy 2539 CTCCTGGAAGATGATTACAGTGTAAACAGAAAGTGAAG 2581
 Db |||||
 Qy 2325 -TTCCTAGAAGGATGCTCAGTGT-GCCGAAGACAGACAAG 2365
 Db |||||

Search completed: June 28, 2003, 06:24:46
 Job time : 5641.03 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:03:00 ; Search time 475.613 Seconds
(without alignments)
15577.944 Million cell updates/sec

Title: US-09-981-353-172
Perfect score: 3290
Sequence: 1 cggcgctctcgtggttc.....agttttgagaaaaaaa 3290

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS2/cgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3270.8	99.4	3328	24	Human DNA sequence
2	3175.2	96.5	3220	24	Human CDNA differe
3	3175.2	96.5	3220	24	Colon adenocarcino
4	3175.2	96.5	3220	24	Breast cancer rela
5	3175.2	96.5	3220	24	Breast cancer rela
6	3173.6	96.5	3220	19	Human sodium-lichi
7	3172	96.4	3211	12	Human gibbon ape 1
8	3172	96.4	3211	16	GalV receptor-enco
9	3172	96.4	3211	20	CDNA encoding the

10	2911.2	88.5	2916	24	ABQ91973	Human NF-kB activa
11	2133	64.8	3227	24	AB199293	Mouse ischaemic co
12	2133	64.8	3260	16	AAQ83680	GLVR-1-encoding CD
13	2131.4	64.8	3260	20	AA15172	DNA encoding the g
14	617.2	18.8	3175	15	AAQ74311	Human amphotropic
15	617.2	18.8	3175	19	AAV33502	Human sodium-lichi
16	617.2	18.8	3175	24	ABL61806	Colon adenocarcino
17	495.2	15.1	662	21	AAZ80539	Human colon cancer
18	459.8	14.0	463	24	ABL79825	Human ovarian canc
19	451.4	13.7	613	21	AAZ80437	Human colon cancer
20	448	13.6	495	24	ABK45753	CDNA encoding colo
21	430	13.1	441	24	ABL79830	Human ovarian canc
22	415	12.6	2747	24	ABL34397	Human immune syste
23	406.4	12.4	414	24	ABL80283	Human ovarian canc
24	396.6	12.1	2747	24	ABL34396	Human immune syste
25	394.4	12.0	396	24	ABL79931	Human ovarian canc
26	393	11.9	643	21	AAZ80185	Human colon cancer
27	324.4	9.9	338	24	ABL86856	Human ovarian canc
28	306.4	9.3	330	24	ABL80745	Human ovarian canc
29	294.8	9.0	372	24	ABQ58882	Human colon cancer
30	285	8.7	285	16	AAT20581	Human gene signatu
31	237.4	7.2	252	21	AAC08438	Human secreted pro
32	236	7.2	535	22	AKS4031	Murine transport a
33	208.6	6.3	2226	23	ABL13171	Drosophila melanog
34	137.2	4.2	287	22	AA124917	Human breast cance
35	136	4.1	6652	23	ABL13170	Drosophila melanog
36	117.6	3.6	6533	24	ABL33708	Human immune syste
37	114	3.5	659	21	AA14433	Aspergillus oryzae
38	114	3.5	34980	22	AAH41226	Pyrococcus abyssi
39	102.4	3.1	289	22	AA16073	Human breast cance
40	99.8	3.0	1263	23	AA53526	Haemophilus influe
41	99.8	3.0	1830121	17	AAT42063	Haemophilus influe
42	93	2.8	1038602	20	AAZ01425	Complete genome se
43	86	2.6	1230025	20	AAZ91990	Nucleotide sequenc
44	84.6	2.6	1269	23	AA531571	Pseudomonas aerugi
45	77	2.3	92407	22	AAF28549	Genomic fragment #

ALIGNMENTS

RESULT 1

ID AAS94856 standard; DNA; 3328 BP.

AC AAS94856;

XX

XX

DT 14-FEB-2002 (first entry)

XX

Human DNA sequence #111 expressed during foam cell differentiation.

DE

XX

Human; foam cell differentiation; atherosclerosis; cerebral stroke;

KW

KW cardiovascular disorder; coronary artery disease; gene therapy; ds.

XX

OS

XX Homo sapiens.

XX

PN WO200177389-A2.

XX

PD 18-OCT-2001.

XX

PF 04-APR-2001; 2001WO-US11128.

XX

PR 05-APR-2000; 2000US-195106P.

XX

XX (INCY-) INCYTE GENOMICS INC.

PA Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T, Tai J;

PI

XX WPI; 2002-010925/01.

DR

XX Composition useful for diagnosis of conditions, disorders or diseases associated with atherosclerosis, comprises several polynucleotides that

Db 1834 CAGACTCCAAGAACGGAATTGGAATGGAAGAGTTACACCAAGTTACTGCAATGCTGTGTCG 1893
Qy 1861 ACCTTCACTCAGCATCTGAGATAGACATGAGTGTCAAGGCAGAGATGGGTCTAGGTGACA 1920
Db 1894 ACCTTCACTCAGCATCTGAGATAGACATGAGTGTCAAGGCAGAGATGGGTCTAGGTGACA 1953
Qy 1921 GAAAGGAAGTAAGTCTCTCTAGAGAAATGGTATGACACAGGATAGACCTGAAAGTCTCTC 1980
Db 1954 GAAAGGAAGTAAGTCTCTCTAGAGAAATGGTATGACACAGGATAGACCTGAAAGTCTCTC 2013
Qy 1981 TCCTCTTCCAGTTCCTGCGAGATCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGGTGGCA 2040
Db 2014 TCCTCTTCCAGTTCCTGCGAGATCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGGTGGCA 2073
Qy 2041 ATG-ACGTAAGCAATGCCATTTGGGCTCTGGTGTCTTTATATTTGGTATGACACAGGA 2099
Db 2074 ATGCAGGTAAGCAATGCCATTTGGGCTCTGGTGTCTTTATATTTGGTATGACACAGGA 2133
Qy 2100 GATGTTTCTTCAAAAGTGGCAACCAATATGGCTTCTACTCTATGGTGGTGTGTATC 2159
Db 2134 GATGTTTCTTCAAAAGTGGCAACCAATATGGCTTCTACTCTATGGTGGTGTGTATC 2193
Qy 2160 TGTGTTGGTCTGTGGGTTTGGGGAAGAAGATTATCCAGACCATGGGGAAGGATCTGACA 2219
Db 2194 TGTGTTGGTCTGTGGGTTTGGGGAAGAAGATTATCCAGACCATGGGGAAGGATCTGACA 2253
Qy 2220 CCGATCACACCTCTAGTGGCTTCACTATTTGAATGGAATGGAATGGAATGGAATGGAAT 2279
Db 2254 CCGATCACACCTCTAGTGGCTTCACTATTTGAATGGAATGGAATGGAATGGAATGGAAT 2313
Qy 2280 GCATCAAAATATTGGCTTCCCATCAGTACAAACATGTAAGTGGGCTCTGTGTGTCT 2339
Db 2314 GCATCAAAATATTGGCTTCCCATCAGTACAAACATGTAAGTGGGCTCTGTGTGTCT 2373
Qy 2340 GTTGGCTGGCTCGGTCCAGAGAGCTGTGACCTGGCTCTTTCGTAACATTTTATG 2399
Db 2374 GTTGGCTGGCTCGGTCCAGAGAGCTGTGACCTGGCTCTTTCGTAACATTTTATG 2433
Qy 2400 GCTGGTGTGTCAAGTCCCATCTTCTGGAGTTATCAGTGTGCAATGCAATGCAATCTTC 2459
Db 2434 GCTGGTGTGTCAAGTCCCATCTTCTGGAGTTATCAGTGTGCAATGCAATGCAATCTTC 2493
Qy 2460 AGATATGATCTCAGATGGAAGCTGTTGAGATTTAAATTTGTGTCAATGTTGGG 2519
Db 2494 AGATATGATCTCAGATGGAAGCTGTTGAGATTTAAATTTGTGTCAATGTTGGG 2553
Qy 2520 ACCATCTTAGGTATTCCTGCTCCCTGAGAGATGATTACAGTGTAAACAGAGACTGACA 2579
Db 2554 ACCATCTTAGGTATTCCTGCTCCCTGAGAGATGATTACAGTGTAAACAGAGACTGACA 2613
Qy 2580 AGAGTCTTTTATTTGGGAGCCAGAGAGGAGTGTACTTGTGTATATACTGCTTTTG 2639
Db 2614 AGAGTCTTTTATTTGGGAGCCAGAGAGGAGTGTACTTGTGTATATACTGCTTTTG 2673
Qy 2640 TGCTAAATATGAATGCTCAAAATAGTGTGTAATAATAGCCGGGTTCCACTGCTCC 2699
Db 2674 TGCTAAATATGAATGCTCAAAATAGTGTGTAATAATAGCCGGGTTCCACTGCTCC 2733
Qy 2700 TGCTAGGTGCTTCTTCTGGGCTGTGAATTCCTGTACATATTTCTACTTTTGT 2759
Db 2734 TGCTAGGTGCTTCTTCTGGGCTGTGAATTCCTGTACATATTTCTACTTTTGT 2793
Qy 2760 ATCAGGCTTCAATTCATATTTTAAATGTTGTCTCTGAAAGATGACTTGTGATTTTTT 2819
Db 2794 ATCAGGCTTCAATTCATATTTTAAATGTTGTCTCTGAGATGACTTGTGATTTTTT 2853
Qy 2820 TTCTTTTTTTTAAACCAAGAGCCGTTTGACAGAGCATGCTCTCGTGTGTGTTTCA 2879
Db 2854 TTCTTTTTTTTAAACCAAGAGCCGTTTGACAGAGCATGCTCTCGTGTGTGTTTCA 2913
Qy 2880 CCAGCTTCTGCTCAGATGCAAGGATTTAAACAAAATATTAACACTACACTCCCT 2939
Db 2914 CCAGCTTCTGCTCAGATGCAAGGATTTAAACAAAATATTAACACTACACTCCCT 2973

Qy 2940 TGTAGTCTCTTATATAAGTAGAGTCTTGGTACTCTGCCCTCTCTGTCAGTAGTGGCAGGA 2999
Db 2974 TGTAGTCTCTTATATAAGTAGAGTCTTGGTACTCTGCCCTCTCTGTCAGTAGTGGCAGGA 3033
Qy 3000 TCTATTGGCATATTCCGGAGCTTCTTAGAGGATGAGGTTCTTTGAACACAGTGAAAAAT 3059
Db 3034 TCTATTGGCATATTCCGGAGCTTCTTAGAGGATGAGGTTCTTTGAACACAGTGAAAAAT 3093
Qy 3060 TAAATTAGTAACCTTTTTCGAAGCAGTTTATTGACTGTTTATTGCTAAAGAAAGTAAGAA 3119
Db 3094 TAAATTAGTAACCTTTTTCGAAGCAGTTTATTGACTGTTTATTGCTAAAGAAAGTAAGAA 3153
Qy 3120 AGAAAAAGCCTCTTGGCAATCTTGGTATTCTTTTAAAGATTCTTGGCAGTGTGGATGGA 3179
Db 3154 AGAAAAAGCCTCTTGGCAATCTTGGTATTCTTTTAAAGATTCTTGGCAGTGTGGATGGA 3213
Qy 3180 TGAATGAAGTGAATGTGAATTTGGGCAAGTTAAATGGGACAGCTTCCATGTTCAATTT 3239
Db 3214 TGAATGAAGTGAATGTGAATTTGGGCAAGTTAAATGGGACAGCTTCCATGTTCAATTT 3273
Qy 3240 GTCTACCTCTTAACCTAAATAAAAAAGCCCTACAGTTTATAGAAAAA 3285
Db 3274 GTCTACCTCTTAACCTAAATAAAAAAGCCCTACAGTTTATAGAAAAA 3319

RESULT 2

ABK83852

ID ABK83852 standard; cDNA; 3220 BP.

XX AC ABK83852;

XX DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #423.

XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
XX granulocyte activation; chronic inflammation; allergy.

OS Homo sapiens.

XX WO200228999-A2.

XX PD 11-APR-2002.

XX 03-OCT-2001; 2001WO-US30821.

XX 03-OCT-2000; 2000US-237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression
XX of genes associated with granulocyte activation, which serves as
XX diagnostic markers that is useful for monitoring disease states and
XX drug toxicity

PS Claim 1; SEQ ID No 423; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing
XX the expression level to an expression level in an unactivated
XX GC, where differential expression of Gs is indicative of GCA.

CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 3220 BP; 783 A; 710 C; 763 G; 964 T; 0 other;

Query Match 96.5%; Score 3175.2; DB 24; Length 3220;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3182; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY	97	CCGACCGGGCGTGTGCGGCGGCTCCAGCGCTCCCGCGCTCGATCTCTCGTCTCCCG	156
DB	24	CCGGGCGGTGCGCGTGTGCGGCGGCTCCAGCGCTCCCGCGCTCGATCTCTCGTCTCCCG	83
QY	157	CTCCGCGCTCCCTTTTCCCTGGATGAATCTTGCGTCTCTCTCTCCGCGCATGGAATTC	216
DB	84	CTCCGCGCTCCCTTTTCCCTGGATGAATCTTGCGTCTCTCTCTCCGCGCATGGAATTC	143
QY	217	TGCTCCGCTGCTTTTACGCTCTCTGAGCAAGAAACCCAGACAAACAGATGCCCATACGC	276
DB	144	TGCTCCGCTGCTTTTACGCTCTCTGAGCAAGAAACCCAGACAAACAGATGCCCATACGC	203
QY	277	AGCGTATAGCAATCTCCCGCTCGGTTCTGTCGCGTAGTTTACAGTATTAATTTT	336
DB	204	AGCGTATAGCAATCTCCCGCTCGGTTCTGTCGCGTAGTTTACAGTATTAATTTT	263
QY	337	ATATAATATATATTTTATATAGCAATTTTATACATCATATCTGTTTACACATCTT	396
DB	264	ATATAATATATATTTTATATAGCAATTTTATACATCATATCTGTTTACACATCTT	323
QY	397	GAAGGCGCTCAGTAGTCTCTTTACTAAACCAACCACTACTCCAGAGAAATGGCAACGGTGA	456
DB	324	GAAGGCGCTCAGTAGTCTCTTTACTAAACCAACCACTACTCCAGAGAAATGGCAACGGTGA	383
QY	457	TTACCACTAGTACTAGCTGTACCGCGCTCTGCTCTGTTGGTAGTACCTATGATGC	516
DB	384	TTACCACTAGTACTAGCTGTACCGCGCTCTGCTCTGTTGGTAGTACCTATGATGC	443
QY	517	TCACTCTGGGCTTCATTATTTGCAATTTGTCTTGCAATTTCTCCGCGGAGGCAATGATGTAG	576
DB	444	TCACTCTGGGCTTCATTATTTGCAATTTGTCTTGCAATTTCTCCGCGGAGGCAATGATGTAG	503
QY	577	CAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTACCTGAAAGCAAGCCTGCATCC	636
DB	504	CAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTACCTGAAAGCAAGCCTGCATCC	563

QY	637	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCCAAAGTGAGCGAAACCA	696
DB	564	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCCAAAGTGAGCGAAACCA	623
QY	697	TCCGGAAGGGCTTGATTGACCTGGAGATGTACAACTCGACTCAAGGGGCTACTGATGGCG	756
DB	624	TCCGGAAGGGCTTGATTGACCTGGAGATGTACAACTCGACTCAAGGGGCTACTGATGGCG	683
QY	757	GCTCAGTCAGTCTATGTTTGGTCTGTGTGGCAACTCTGTCCTCTGGTTCGTTTGAAGC	816
DB	684	GCTCAGTCAGTCTATGTTTGGTCTGTGTGGCAACTCTGTCCTCTGGTTCGTTTGAAGC	743
QY	817	TCCCTATTCTTGGAAACCCATTGTTTGGTGGCAACTATTGGTTCCTCCCTCGTGGCAA	876
DB	744	TCCCTATTCTTGGAAACCCATTGTTTGGTGGCAACTATTGGTTCCTCCCTCGTGGCAA	803
QY	877	AGGGCAGGAGGGTGTCAAGTGGTCTGAACTGATAAATAATTTGATGTCCTGGTTCGTGT	936
DB	804	AGGGCAGGAGGGTGTCAAGTGGTCTGAACTGATAAATAATTTGATGTCCTGGTTCGTGT	863
QY	937	CCCACTGCTTCTTGGAAATTTATGCTGAAATTTTATCTTCTGCTGGTTCGTCATTCATCC	996
DB	864	CCCACTGCTTCTTGGAAATTTATGCTGAAATTTTATCTTCTGCTGGTTCGTCATTCATCC	923
QY	997	TCCATAAGGCAGATCCAGTTCCTTAATGTTTGGAGCTTTGCCAGTTCCTATGCTGCA	1056
DB	924	TCCATAAGGCAGATCCAGTTCCTTAATGTTTGGAGCTTTGCCAGTTCCTATGCTGCA	983
QY	1057	CAGTTGGAAATAAACCTCTTTTCCATCATGTATACCTGAGACACCGTTCCTGGCTTTGACA	1116
DB	984	CAGTTGGAAATAAACCTCTTTTCCATCATGTATACCTGAGACACCGTTCCTGGCTTTGACA	1043
QY	1117	AACCTCTCTGTGGGTACCATCTCTCATCTCGTGGGATGTCAGTTCCTTCTGTCCTTTA	1176
DB	1044	AACCTCTCTGTGGGTACCATCTCTCATCTCGTGGGATGTCAGTTCCTTCTGTCCTTTA	1103
QY	1177	TCGCTGCTGTTCTTTGTATGTCCAGGATGAAGAGAAAAATTTGAACAGAAAAATAAGTGA	1236
DB	1104	TCGCTGCTGTTCTTTGTATGTCCAGGATGAAGAGAAAAATTTGAACAGAAAAATAAGTGA	1163
QY	1237	GTCTCTTCTGAAGCCCTTAAATGAAAAAAGAAATAGCTTGAAGAAAGCAATGAAGAAA	1296
DB	1164	GTCTCTTCTGAAGCCCTTAAATGAAAAAAGAAATAGCTTGAAGAAAGCAATGAAGAAA	1223
QY	1297	CAAGTTCCTGTTGTTGATATTTGAAACAAAGCATCTCTGTTCTGAGGTAGGGCTGCA	1356
DB	1224	CAAGTTCCTGTTGTTGATATTTGAAACAAAGCATCTCTGTTCTGAGGTAGGGCTGCA	1283
QY	1357	CTGTGCCCCCTCCAGGCTGTGGTGGAGAGAGAAACAGTCTCATTTCAAACTTTGAGATTTGG	1416
DB	1284	CTGTGCCCCCTCCAGGCTGTGGTGGAGAGAGAAACAGTCTCATTTCAAACTTTGAGATTTGG	1343
QY	1417	AGGAAGCTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGAGAAACACCATAG	1476
DB	1344	AGGAAGCTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGAGAAACACCATAG	1403
QY	1477	ATAGCAGCGTGAATGGTGCAGTGCAGTTCCCTTAATGGGAACTTCTGTCAGTTTCAGTCAAG	1536
DB	1404	ATAGCAGCGTGAATGGTGCAGTGCAGTTCCCTTAATGGGAACTTCTGTCAGTTTCAGTCAAG	1463
QY	1537	CCGTGAGCAACCAAAATAAATCCAGTGGCACTACAGTATCACACCGTGCATTAAGGATT	1596
DB	1464	CCGTGAGCAACCAAAATAAATCCAGTGGCACTACAGTATCACACCGTGCATTAAGGATT	1523
QY	1597	CCGGCTCTGTACAAAGAGCTTCTCCATAAATTAATCATCTTTGCCAAGGTGGAGATTCATGG	1656
DB	1524	CCGGCTCTGTACAAAGAGCTTCTCCATAAATTAATCATCTTTGCCAAGGTGGAGATTCATGG	1583
QY	1657	GAGACTCCGGTGACAAACCTTAAGGCGCAATAATAGCTATATCTTCTATACCATGGCAA	1716
DB	1584	GAGACTCCGGTGACAAACCTTAAGGCGCAATAATAGCTATATCTTCTATACCATGGCAA	1643

1297 CAAAGTTGCTCTGTTGGTGATATTGAAAACAAGCATCCTGTTCTGAGGTAGGCCCTGGCCA 1356
1224 CAAAGTTGCTGTTGGTGATATTGAAAACAAGCATCCTGTTCTGAGGTAGGCCCTGGCCA 1283
1357 CTGTGCCCTCCAGGCTGTGGTGAGAGAGAACAGTCTCATTCAAACCTTTGGAGATTTGG 1416
1284 CTGTGCCCTCCAGGCTGTGGTGAGAGAGAACAGTCTCATTCAAACCTTTGGAGATTTGG 1343
1417 AGGAAGCTCAGAGAGAGAGAGCTTCCAGCGTGCACTTTGAAAGAGGAAACAGCATAG 1476
1344 AGGAAGCTCAGAGAGAGAGAGCTTCCAGCGTGCACTTTGAAAGAGGAAACAGCATAG 1403
1477 ATAGCACCGTGAATGCTGAGTGCAGTTGCCCTTAATGGGAACCTTGTCCAGTTCACTCAAG 1536
1404 ATAGCACCGTGAATGCTGAGTGCAGTTGCCCTTAATGGGAACCTTGTCCAGTTCACTCAAG 1463
1537 CCGTCAAGCAACCAATAAATCCAGTGGCCACTACCAGTATCACACCGTGCAATAGGATT 1596
1464 CCGTCAAGCAACCAATAAATCCAGTGGCCACTACCAGTATCACACCGTGCAATAGGATT 1523
1597 CCGGCTGTACAAAGAGCTACTCCATAAATTACATCTTCCAAAGTGGAGATTGCATGG 1656
1524 CCGGCTGTACAAAGAGCTACTCCATAAATTACATCTTCCAAAGTGGAGATTGCATGG 1583
1657 GAGACTCCGCTGACAAACCCCTTAAGCGGCAATAATAGCTATACTTCCATACATGGCAA 1716
1584 GAGACTCCGCTGACAAACCCCTTAAGCGGCAATAATAGCTATACTTCCATACATGGCAA 1643
1717 TATGTGGCATGCTCTGGGATTCAATCCGTGCCAAAGAGGTGAACAGAGGGCGAAGAA 1776
1644 TATGTGGCATGCTCTGGGATTCAATCCGTGCCAAAGAGGTGAACAGAGGGCGAAGAA 1703
1777 TGGAGAAGCTGACATGGCCCTAATGCAGACTCCAAGAGGGAATTCGAATGGACAGTTACA 1836
1704 TGGAGAAGCTGACATGGCCCTAATGCAGACTCCAAGAGGGAATTCGAATGGACAGTTACA 1763
1837 CCAGTTACTGCAATGCTGTCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGCA 1896
1764 CCAGTTACTGCAATGCTGTCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGCA 1823
1897 AGGCAGAGATGGTCTAGGTGACAGAAAAGGAAGTAATGGCTCTCTAGAAAGATGCTATG 1956
1824 AGGCAGAGATGGTCTAGGTGACAGAAAAGGAAGTAATGGCTCTCTAGAAAGATGCTATG 1883
1957 ACCAGGATAAGCTGTAAGTCTCTCTCTCTCTCCAGTTCCTGAGATCCTTACAGCCCTGCT 2016
1884 ACCAGGATAAGCTGTAAGTCTCTCTCTCTCTCCAGTTCCTGAGATCCTTACAGCCCTGCT 1943
2017 TTGGGTCATTCCGCCATGGTGGCAATGACGTAAAGCAATGCCATTTGGGCTCTGTTGCTT 2076
1944 TTGGGTCATTCCGCCATGGTGGCAATGACGTAAAGCAATGCCATTTGGGCTCTGTTGCTT 2003
2077 TATATTTGGTTTATGACACAGAGATGTTTCTTCAAAGTGCAACACCAATATGGCTTC 2136
2004 TATATTTGGTTTATGACACAGAGATGTTTCTTCAAAGTGCAACACCAATATGGCTTC 2063
2137 TACTCTATGGTGGTGTGATCTGTGTTGGTCTGTGGGTTGGGGAAGAAGAGTTATCC 2196
2064 TACTCTATGGTGGTGTGATCTGTGTTGGTCTGTGGGTTGGGGAAGAAGAGTTATCC 2123
2197 AGACCATGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTCACTATTAAGCTGG 2256
2124 AGACCATGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTCACTATTAAGCTGG 2183
2257 CATCTGCCCTCACTGTTGGTGAATGCAATTAATTTGGCCTTCCCATCAGTACACACATT 2316
2184 CATCTGCCCTCACTGTTGGTGAATGCAATTAATTTGGCCTTCCCATCAGTACACACATT 2243
2317 GTAAAGTGGGCTCTGTTGTCTGTGTTGGCTCCGGTCCCAAGAGGCTGTTGACTGGC 2376
2244 GTAAAGTGGGCTCTGTTGTCTGTGTTGGCTCCGGTCCCAAGAGGCTGTTGACTGGC 2303
2377 GTCTCTTTTCGTAACATTTTATATGGCCTGTTTGTACAGTCCCATTTCTCGAGTTATCA 2436

RESULT 4

ABU62867

ID ABL62867 standard; DNA; 3220 BP.

XX ABL62867;

XX AC

DT 15-MAY-2002 (first entry)

XX

DE Breast cancer related gene sequence SEQ ID NO:1204.

2304 GTCTCTTTTCGTAACATTTTATGGCTCGTGTGTGCACAGTCCCATTTCTCGAGTTATCA 2363
2437 GTGCTGCCATCATATGCAATCTTCAGATATGTCATCCTCAGAATGTGAAGCTGTTTGAGAT 2496
2364 GTGCTGCCATCATATGCAATCTTCAGATATGTCATCCTCAGAATGTGAAGCTGTTTGAGAT 2423
2497 TAAAAATTTGTGTCAATGTTTGGGACCACTCTTAAAGTATTCTCTCCCTCCCTGAAGAATGATT 2556
2424 TAAAAATTTGTGTCAATGTTTGGGACCACTCTTAAAGTATTCTCTCCCTCCCTGAAGAATGATT 2483
2557 ACAGTGTAAACAGAGACTGCAAGAGTCTTTTATTTTGGGAGCCAGAGGAGGAAGTGT 2616
2484 ACAGTGTAAACAGAGACTGCAAGAGTCTTTTATTTTGGGAG-CAGAGGAGGAAGTGT 2542
2617 TACTTGTGCTATAAATGCTGCTAAATATGAATGTCTCAAAATTAGCTGTCTGAAA 2676
2543 TACTTGTGCTATAAATGCTGCTAAATATGAATGTCTCAAAATTAGCTGTCTGAAA 2602
2677 ATAGCCCGGTTCCACTGGCTCCTGCTGAGGTCCCTTTCTCTGGGCTGTGAATTCCT 2736
2603 ATAGCCCGGTTCCACTGGCTCCTGCTGAGGTCCCTTTCTCTGGGCTGTGAATTCCT 2662
2737 GTACATATTTCTCTACTTTTGTATCAGGCTTCAATTCATTTATGTTTAAATGTTGCTC 2796
2663 GTACATATTTCTCTACTTTTGTATCAGGCTTCAATTCATTTATGTTTAAATGTTGCTC 2722
2797 TGAAGTACATCTGTGATTTTCTTTTAAACCATGAAGAGCCGTTTGCAGAG 2856
2723 TGAAGTACATCTGTGATTTTCTTTTAAACCATGAAGAGCCGTTTGCAGAG 2782
2857 CATGCTGCGTGTGTGGTTTCCAGAGCTTCTGCCCTCACATGCAAGGATTTAAACAAC 2916
2783 CATGCTGCGTGTGTGGTTTCCAGAGCTTCTGCCCTCACATGCAAGGATTTAAACAAC 2842
2917 AAAAATATACTAACAATTCCTTGTAGTCTCTTATATAGTAGAGTCTTGGTACTCTG 2976
2843 AAAAATATACTAACAATTCCTTGTAGTCTCTTATATAGTAGAGTCTTGGTACTCTG 2902
2977 CCCTCTGTGCTAGTGGCAGGATCTATTGGCATATTCGGGAGCTTCTTAGAGGATGAG 3036
2903 CCCTCTGTGCTAGTGGCAGGATCTATTGGCATATTCGGGAGCTTCTTAGAGGATGAG 2962
3037 GTTCTTTGAACAACAGTGAATAATTTAAATTAGTAACTTTTTCGAAAGCAGTTTATTGACTG 3096
2963 GTTCTTTGAACAACAGTGAATAATTTAAATTAGTAACTTTTTCGAAAGCAGTTTATTGACTG 3022
3097 TTATTTGCTTAAGAGAGTAAGAAAGAAAGAGCTTGTGGCAATCTTGGTTATTCTTTAA 3156
3023 TTATTTGCTTAAGAGAGTAAGAAAGAAAGAGCTTGTGGCAATCTTGGTTATTCTTTAA 3082
3157 GATTTCTGGCAGTGTGGGATGGAATGAAGTGAATGTGAATTTGGGCAAGTTAAAT 3216
3083 GATTTCTGGCAGTGTGGGATGGAATGAAGTGAATGTGAATTTGGGCAAGTTAAAT 3142
3217 GGGACAGCTTCCATGTTTCATTGTTCTACCTCTTAACCTGAATAAAAAAGCCTTACAGTTT 3276
3143 GGGACAGCTTCCATGTTTCATTGTTCTACCTCTTAACCTGAATAAAAAAGCCTTACAGTTT 3202
3277 TAGAAAAA 3285
3203 TAGAAAAA 3211

RESULT 4

ABU62867

ID ABL62867 standard; DNA; 3220 BP.

XX ABL62867;

XX AC

DT 15-MAY-2002 (first entry)

XX

DE Breast cancer related gene sequence SEQ ID NO:1204.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; db.
OS Homo sapiens.
XX WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US10838.
XX 05-JUN-2000; 2000US-209473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233617P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234034P.
XX 20-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234924P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 25-SEP-2000; 2000US-235280P.
XX 26-SEP-2000; 2000US-235637P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.
XX 28-SEP-2000; 2000US-236028P.
XX 28-SEP-2000; 2000US-236032P.
XX 28-SEP-2000; 2000US-236033P.
XX 28-SEP-2000; 2000US-236034P.
XX 28-SEP-2000; 2000US-236109P.
XX 28-SEP-2000; 2000US-236111P.
XX 29-SEP-2000; 2000US-236842P.
XX 02-OCT-2000; 2000US-237172P.
XX 02-OCT-2000; 2000US-237173P.
XX 02-OCT-2000; 2000US-237278P.
XX 02-OCT-2000; 2000US-237294P.
XX 02-OCT-2000; 2000US-237295P.
XX 02-OCT-2000; 2000US-237316P.
XX 03-OCT-2000; 2000US-237425P.
XX 03-OCT-2000; 2000US-237588P.
XX 03-OCT-2000; 2000US-237604P.
XX 03-OCT-2000; 2000US-237606P.
XX 03-OCT-2000; 2000US-237608P.
XX 01-NOV-2000; 2000US-244867P.
XX 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
PS Claim 1; SEQ ID 1204; 44pp; English.
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in

CC expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB161664 to ABL70110), or is at least 95% identical to (S), where a change in CC expression is indicative of anti-neoplastic activity. (I) has cytostatic CC activity and can be used in gene therapy. M1 can be used for screening CC an anti-neoplastic agent, and can be used for producing a product which CC is the data collected with respect to the anti-neoplastic agent as a CC result of M1, and the data is sufficient to convey the chemical CC structure and/or properties of the agent. M1 can be used in the CC treatment of cancer such as colon, breast, stomach, lung, thyroid, CC oesophageal, ovarian, kidney, prostate or pancreatic cancer, CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 3220 BP; 783 A; 710 C; 763 G; 964 T; 0 other;
Query Match 96.5%; Score 3175.2; DB 24; Length 3220;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3182; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 97 CGACCCGGCGGTGCGCGTCCAGCGCTCCAGCGCTCCGATCTCTCTCCGTCCTCCG 156
DB 24 CCGGGCGGTGCGCGTCCAGCGCTCCAGCGCTCCGATCTCTCTCCGTCCTCCG 83
QY 157 CTCGCCCTCCCTTTTCCCTGGATGAACCTTGGCTCTTCTCTCTCCGCCATGGAATTC 216
DB 84 CTCGCCCTCCCTTTTCCCTGGATGAACCTTGGCTCTTCTCTCTCCGCCATGGAATTC 143
QY 217 TGCTCCGTGTTTGTAGCCCTCTGAGCCAAAGAAAACCCAGACAAACAGATGCCATAGCC 276
DB 144 TGCTCCGTGTTTGTAGCCCTCTGAGCCAAAGAAAACCCAGACAAACAGATGCCATAGCC 203
QY 277 AGCGTATAGCAGTAACCTCCCGCTCTGTCGCGTAGTTTACAGTATTAATTTT 336
DB 204 AGCGTATAGCAGTAACCTCCCGCTCTGTCGCGTAGTTTACAGTATTAATTTT 263
QY 337 ATATAAT 396
DB 264 ATATAAT 323
QY 397 GAAAGCGCTCAGTAGTTCTCTTACTAAACCACTACTCCAGAGAAATGGCAAGCTGA 456
DB 324 GAAAGCGCTCAGTAGTTCTCTTACTAAACCACTACTCCAGAGAAATGGCAAGCTGA 383
QY 457 TTACCACTACTACAGCTGCTACCGCGCTCTGGTCTCTTGGTGACCTATCGATGC 516
DB 384 TTACCACTACTACAGCTGCTACCGCGCTCTGGTCTCTTGGTGACCTATCGATGC 443
QY 517 TCATCTCGGCTTCATTATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCAT 576
DB 444 TCATCTCGGCTTCATTATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCAT 503
QY 577 CAAATTTTGGTACAGCTGCGCTCAGGTGTAGTACCTGACCTGAAGCAAGCTGCATCC 636
DB 504 CAAATTTTGGTACAGCTGCGCTCAGGTGTAGTACCTGACCTGAAGCAAGCTGCATCC 563
QY 637 TAGCTAGCATCTTTGAAACAGTGGGCTCTGCTTACTGGGGCCAAAGTGAGGAAACCA 696
DB 564 TAGCTAGCATCTTTGAAACAGTGGGCTCTGCTTACTGGGGCCAAAGTGAGGAAACCA 623
QY 697 TCCGGAAGGGCTTGATTGACGTGGAGATGTACAACCTCGACTCAAGGGCTACTGTAGGCGG 756
DB 624 TCCGGAAGGGCTTGATTGACGTGGAGATGTACAACCTCGACTCAAGGGCTACTGTAGGCGG 683
QY 757 GCTCAGTCAGTCTAGTTGGTTCTGCTGTGTGCAACTCTGTGCTGCTGCTGCTGCTGCTG 816
DB 684 GCTCAGTCAGTCTAGTTGGTTCTGCTGTGTGCAACTCTGTGCTGCTGCTGCTGCTGCTG 743
QY 817 TCCCTATTTCTGGAACCCANTTGTATTTGTTGGTGAACCTATTGCTTCTCCCTCGTGCA 876
DB 744 TCCCTATTTCTGGAACCCANTTGTATTTGTTGGTGAACCTATTGCTTCTCCCTCGTGCA 803

[illegible]

Db	2863	GTTCCTTTGAACACAGTGAATAATTTAAATTAGTAACCTTTTTTGGCAAGCAGTTTATTGACTG	3022
Qy	3097	TTATTGCTTAAGAAAGAAGTAAAGAAAAGCCCTGTTGGCAACTCTTGCTGTTATTTCTTTAA	3156
Db	3023	TTATTGCTTAAGAAAGAAGTAAAGAAAAGCCCTGTTGGCAACTCTTGCTGTTATTTCTTTAA	3082
Qy	3157	GATTTCTGGCAGTGTGGATGGATGAATGAAGTGGAAATGTGAACCTTTGGCAGAGTTAAAT	3216
Db	3083	GATTTCTGGCAGTGTGGATGGATGAATGAAGTGGAAATGTGAACCTTTGGCAGAGTTAAAT	3142
Qy	3217	GGGACAGCCTTCAGATGTTTCATTTCTCTACCTCTTTAACTGAATAAAAAAGCCTACAGTTTT	3276
Db	3143	GGGACAGCCTTCAGATGTTTCATTTCTCTACCTCTTTAACTGAATAAAAAAGCCTACAGTTTT	3202
Qy	3277	TAGAAAAA	3285
Db	3203	TAGAAAAA	3211

RESULT 5	
ABL63086	
ID	ABL63086 standard; DNA; 3220 BP.
XX	
XX	
AC	ABL63086;
XX	
XX	
DT	15-MAY-2002 (first entry)
XX	
XX	
DE	Breast cancer related gene sequence SEQ ID NO:1423.
XX	
XX	
KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW	gene; ds.
XX	
OS	Homo sapiens.
OS	
XX	
PN	WO200194629-A2.
XX	
XX	
PD	13-DEC-2001.

PR	05-JUN-2000	2000US-209473P.
PR	05-JUN-2000	2000US-209531P.
PR	18-SEP-2000	2000US-233133P.
PR	18-SEP-2000	2000US-233617P.
PR	20-SEP-2000	2000US-233400P.
PR	20-SEP-2000	2000US-234034P.
PR	20-SEP-2000	2000US-234052P.
PR	22-SEP-2000	2000US-234509P.
PR	22-SEP-2000	2000US-234567P.
PR	25-SEP-2000	2000US-234923P.
PR	25-SEP-2000	2000US-234924P.
PR	25-SEP-2000	2000US-235077P.
PR	25-SEP-2000	2000US-235082P.
PR	25-SEP-2000	2000US-235134P.
PR	25-SEP-2000	2000US-235280P.
PR	26-SEP-2000	2000US-235637P.
PR	26-SEP-2000	2000US-235638P.
PR	27-SEP-2000	2000US-235711P.
PR	27-SEP-2000	2000US-235720P.
PR	27-SEP-2000	2000US-235840P.
PR	27-SEP-2000	2000US-235863P.
PR	28-SEP-2000	2000US-236028P.
PR	28-SEP-2000	2000US-236032P.
PR	28-SEP-2000	2000US-236033P.
PR	28-SEP-2000	2000US-236034P.
PR	28-SEP-2000	2000US-236109P.
PR	28-SEP-2000	2000US-236111P.
PR	29-SEP-2000	2000US-236842P.
PR	29-SEP-2000	2000US-236891P.
PR	02-OCT-2000	2000US-237172P.
PR	02-OCT-2000	2000US-237173P.

PR	02-OCT-2000; 2000US-237278P.
PR	02-OCT-2000; 2000US-237294P.
PR	02-OCT-2000; 2000US-237295P.
PR	02-OCT-2000; 2000US-237316P.
PR	03-OCT-2000; 2000US-237425P.
PR	03-OCT-2000; 2000US-237598P.
PR	03-OCT-2000; 2000US-237604P.
PR	03-OCT-2000; 2000US-237606P.
PR	03-OCT-2000; 2000US-237608P.
PR	01-NOV-2000; 2000US-244867P.
PR	01-NOV-2000; 2000US-245084P.
XX	(AVAL-) AVALON PHARM.
XX	
PI	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI	Soppet DR, Weaver Z;
XX	
DR	WPI; 2002-198264/24.
XX	
PT	Screening for anti-neoplastic agent involves exposing cells to a
PT	chemical agent to be tested for anti-neoplastic activity, and
PT	determining a change in expression of a gene of a signature gene set -
XX	
PS	Claim 1; SEQ ID 1423; 44pp; English.
XX	
CC	The present invention describes a method (M1) for screening for an
CC	anti-neoplastic agent. The method involves exposing cells to a chemical
CC	agent to be tested for anti-neoplastic activity, determining a change in
CC	expression of at least one gene (I) of a signature gene set, where (I)
CC	comprises a sequence (S) selected from 8447 sequences (given in ABU61664
CC	to ABU70110), or is at least 95% identical to (S), where a change in
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC	activity and can be used in gene therapy. M1 can be used for screening
CC	an anti-neoplastic agent, and can be used for producing a product which
CC	is the data collected with respect to the anti-neoplastic agent as a
CC	result of M1, and the data is sufficient to convey the chemical
CC	structure and/or properties of the agent. M1 can be used in the
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC	oesophagal, ovarian, kidney, prostate or pancreatic cancer,
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC	carcinoma, papillary carcinoma and Wilms' tumour.
XX	
SQ	Sequence 3220 BP; 783 A; 710 C; 763 G; 964 T; 0 other;
Query Match 96.5%; Score 3175.2; DB 24; Length 3220;	
Best Local Similarity 99.8%; Pred. No. 0;	
Matches 3182; Conservative 0; Mismatches 6; Indels 1; Gaps 1;	
Qy	97 CCGACCCGGGCGGTGCCTCGTCCAGCGGTGCCCGCTCGATCTCCTCGTCCTCCG 156
Dd	
Dd	24 CCAGCGCGTGCCCGTGCCTCGAGCGGTGCCCGCTCGATCTCCTCGTCCTCCG 83
Qy	157 CTCGGCCCTCCCTTTCCCTGGATGAATTCGGTCTCTTCTTCGCGCATGGAATTC 216
Dd	84 CTCGGCCCTCCCTTTCCCTGGATGAATTCGGTCTCTTCTTCGCGCATGGAATTC 143
Qy	217 TGCTCCGTGCTTTTAGCCCTCTCGAGCAAAAGAACCCAGACAACAGATGCCCATAGC 276
Dd	144 TGCTCCGTGCTTTTAGCCCTCTCGAGCAAAAGAACCCAGACAACAGATGCCCATAGC 203
Qy	277 AGCGTATAGCAGTAACCTCCCNGCTCGGTTTTCTGCGCGTAGTTTACAGTATTTTATTTT 336
Dd	204 AGCGTATAGCAGTAACCTCCCNGCTCGGTTTTCTGCGCGTAGTTTACAGTATTTTATTTT 263
Qy	337 ATATAATATATATTATTATATAGCATTTTTTGATACCTCATATTCTGTTTACACATCTT 396
Dd	264 ATATAATATATATTATTATATAGCATTTTTTGATACCTCATATTCTGTTTACACATCTT 323
Qy	397 GAAGAGCGCTCAGTAGTTCTCTTACTAAACCACTACTCCAGAGAATGGCAACGGTGA 456
Dd	324 GAAGAGCGCTCAGTAGTTCTCTTACTAAACCACTACTCCAGAGAATGGCAACGGTGA 383

Qy	457	TTACCAAGTACTACAGCTGCTACCGCCGGCTTCCTGGTGCTCTTTGGTGGACCTACCTATGGATGC	516
Db	384	TTACCAGTACTACAGCTGCTACCGCCGGCTTCCTGGTGCTCTTTGGTGGACCTACCTATGGATGC	443
Qy	517	TCATCTGGGGCTTCATTATTTCATTTGTCTTTGGCATTTCTCCGTGGGAGCCAATGATGAG	576
Db	444	TCATCTGGGGCTTCATTATTTCATTTGTCTTTGGCATTTCTCCGTGGGAGCCAATGATGAG	503
Qy	577	CAAAATCTTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCCCTGAAGCAAAGCTGCATCC	636
Db	504	CAAAATCTTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCCCTGAAGCAAAGCTGCATCC	563
Qy	637	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTCTGGGGGCCAAGAAGTAGCGAAAACCA	696
Db	564	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTCTGGGGGCCAAGAAGTAGCGAAAACCA	623
Qy	697	TCCGGAAGGGCTTCATTAGACGTGGAGATGTACAACCTCGACCTCAAGGGCTACTGATGSCCG	756
Db	624	TCCGGAAGGGCTTCATTAGACGTGGAGATGTACAACCTCGACCTCAAGGGCTACTGATGSCCG	683
Qy	757	GCTCAGTCAGTCTATGTTTGGTTCTGCTGTGTGGCAAACCTCGTGGCTTCGTTTTTGAAGC	816
Db	684	GCTCAGTCAGTCTATGTTTGGTTCTGCTGTGTGGCAAACCTCGTGGCTTCGTTTTTGAAGC	743
Qy	817	TCCCTATTCTGGAACCCATTGTATTGTGTGGTGCAACTATTTGGTTTTCTCCCTCGTGCAAA	876
Db	744	TCCCTATTCTGGAACCCATTGTATTGTGTGGTGCAACTATTTGGTTTTCTCCCTCGTGCAAA	803
Qy	877	AGGGCAGGAGGGTGCAAGTGTCTGAACTGATTAATAATTTGTCATGCTCTTGGTTCTGTGT	936
Db	804	AGGGCAGGAGGGTGCAAGTGTCTGAACTGATTAATAATTTGTCATGCTCTTGGTTCTGTGT	863
Qy	937	CCCCACTGCTTTCTGGGAATTATGTCTGGAATTTTATTTCTTCTCGTTCGTGCATTTCATCC	996
Db	864	CCCCACTGCTTTCTGGGAATTATGTCTGGAATTTTATTTCTTCTCGTTCGTGCATTTCATCC	923
Qy	997	TCCATAAGGCAGATCCAGTTCCTAATGGTTTGCAGCTTTTGCCAGTTCCTATGCCCTGCA	1056
Db	924	TCCATAAGGCAGATCCAGTTCCTAATGGTTTGCAGCTTTTGCCAGTTCCTATGCCCTGCA	983
Qy	1057	CAGTTTGAATAAACCTCTTTTCCATCATGTATCTCGGAGCACCGTCTCGGGCTTTTGACA	1116
Db	984	CAGTTTGAATAAACCTCTTTTCCATCATGTATCTCGGAGCACCGTCTCGGGCTTTTGACA	1043
Qy	1117	AAC TTC TCT GTG GGG T ACC AT C CT CAT CTC CGT GGG ATG TGC AGT TTT CTG TG CCCT TA	1176
Db	1044	AAC TTC TCT GTG GGG T ACC AT C CT CAT CTC CGT GGG ATG TGC AGT TTT CTG TG CCCT TA	1103
Qy	1177	TCGTCCTGGTTCTTTGTATGTCCCAGGATGAAGAGAAAAATTTGAACGAGAAATAAAGTGTA	1236
Db	1104	TCGTCCTGGTTCTTTGTATGTCCCAGGATGAAGAGAAAAATTTGAACGAGAAATAAAGTGTA	1163
Qy	1237	GTCCCTTCGAAAGCCCCCTTAATGAAAAAAAAGAAATAGCTTTGAAAGAGAACCATGAAGAAA	1296
Db	1164	GTCCCTTCGAAAGCCCCCTTAATGAAAAAAAAGAAATAGCTTTGAAAGAGAACCATGAAGAAA	1223
Qy	1297	CAAAAGTTGTCTGTTGGTGATATTCAAACCAAGCATCTCTGTTTCTGAGGTAGGGCCCTGCCA	1356
Db	1224	CAAAAGTTGTCTGTTGGTGATATTGAAAAACAAGCATCTCTGTTTCTGAGGTAGGGCCCTGCCA	1283
Qy	1357	CTGTGCCCCCTCCAGGCTGTGTGGAGAGAGAACAGTCTCATTTCAAACTTTGGAGATTTGG	1416
Db	1284	CTGTGCCCCCTCCAGGCTGTGTGGAGAGAGAACAGTCTCATTTCAAACTTTGGAGATTTGG	1343
Qy	1417	AGGAAGCTCCNAGACAGAGAGAGGCTTCCACGCTGGACTTTGAAAGAGAGAAACGAGCATAG	1476
Db	1344	AGGAAGCTCCNAGACAGAGAGAGGCTTCCACGCTGGACTTTGAAAGAGAGAAACGAGCATAG	1403
Qy	1477	ATAGCACCGTGAATGGTGCACTGAGTTGCCCTAATGSGAACCTTGTCCAGTTCAGTCAAG	1536
Db	1404	ATAGCACCGTGAATGGTGCACTGAGTTGCCCTAATGSGAACCTTGTCCAGTTCAGTCAAG	1463
Qy	1537	CCGTGAGCAACAAATAAATCCAGTGGCCCACTACCAAGTATCACACCGTGCATAGGATT	1596

[illegible]

Db 2543 TACTTGCTGCTATAAAGCTTTTGTGCTAAATATGAATTTCTCAAAATTTAGCTGTGTAA 2602
Qy 2677 ATAGCCGGGTTCCACTGCTCTGCTGAGTCCCTTTCTCTGCGCTGTGAATTCCT 2736
Db 2603 ATAGCCGGGTTCCACTGCTCTGCTGAGTCCCTTTCTCTGCGCTGTGAATTCCT 2662
Qy 2737 GTACATATTTCTCTACTCTTTTGTATCAGGCTTCAATTTCCATTTATGTTTAAATGTTGTC 2796
Db 2663 GTACATATTTCTCTACTCTTTTGTATCAGGCTTCAATTTCCATTTATGTTTAAATGTTGTC 2722
Qy 2797 TGAAGATGACTGTGATTTTTTTTTTTTTTTTTTTTAAACCATGAAGCCGTTTGACAGAG 2856
Db 2723 TGAAGATGACTGTGATTTTTTTTTTTTTTTTTTTTAAACCATGAAGCCGTTTGACAGAG 2782
Qy 2857 CATGCTCTGGTGTGTTTGTCTTACCAGCTTCTGCCCTCATGTCACAGGATTTAACAC 2916
Db 2783 CATGCTCTGGTGTGTTTGTCTTACCAGCTTCTGCCCTCATGTCACAGGATTTAACAC 2842
Qy 2917 AAAAATATAACTACAATTTCCCTTTGAGTCTCTTATATAAGTAGAGTCTTGTGACTCTG 2976
Db 2843 AAAAATATAACTACAATTTCCCTTTGAGTCTCTTATATAAGTAGAGTCTTGTGACTCTG 2902
Qy 2977 CCTCTCTGTCAGTAGTGGCAGGATCTATTGGCATATTCGGAGCTTTCTTAGAGGGATGAG 3036
Db 2903 CCTCTCTGTCAGTAGTGGCAGGATCTATTGGCATATTCGGAGCTTTCTTAGAGGGATGAG 2962
Qy 3037 GTTCTTTGAACAGTGAATTTAAATTTAGTAACTTTTTCGAACAGTTTTATTGACTG 3096
Db 2963 GTTCTTTGAACAGTGAATTTAAATTTAGTAACTTTTTCGAACAGTTTTATTGACTG 3022
Qy 3097 TTATTCCTAAGAGAGTAGTGAAGAAAGAAAGCTGTTGGCAATCTTGGTATTCTTTTAA 3156
Db 3023 TTATTCCTAAGAGAGTAGTGAAGAAAGAAAGCTGTTGGCAATCTTGGTATTCTTTTAA 3082
Qy 3157 GATTTCTGCGAGTGTGGGATGGAATGAAGTGAATGTGAATTTTGGGCAAGTTAAAT 3216
Db 3083 GATTTCTGCGAGTGTGGGATGGAATGAAGTGAATGTGAATTTTGGGCAAGTTAAAT 3142
Qy 3217 GGGACAGCTTCCATGTTCAATTTGTCTACCTCTTAACCTGAATTAAGAAAGCTTACAGTTT 3276
Db 3143 GGGACAGCTTCCATGTTCAATTTGTCTACCTCTTAACCTGAATTAAGAAAGCTTACAGTTT 3202
Qy 3277 TAGAAAAA 3285
Db 3203 TAGAAAAA 3211

RESULT 6

AAV33501

ID AAV33501 standard; cDNA; 3220 BP.

XX AC AAV33501;

XX 18-JAN-1999 (first entry)

XX Human sodium-lithium countertransporter hPiT-1 cDNA.

XX Sodium-lithium countertransporter; sodium-phosphate cotransporter;

XX PiT-1; human; leukaemia virus receptor 1; lithium therapy;

XX manic depression; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 371..2410

XX FT /*tag= a

XX FT /product= leukaemia virus receptor 1

XX W09838203-A1.

XX PD 03-SEP-1998.

XX PF 11-FEB-1998; 98MO-US02875.

XX 27-FEB-1997; 97US-0039462.
PR (UYEM-) UNIV EMORY.
XX Gunn RB, Timmer RT;
PI WPI: 1998-520759/44.
XX P-PSDB; AAW70498.
PT New isolated lithium-sodium counter-transporter DNA - used to
PT develop products for evaluating lithium-sodium transport in
PT erythrocytes, particularly for lithium therapy in manic depression.
XX Claim 3; Page 35-37; 64pp; English.
CC This DNA molecule encodes human lithium-sodium countertransporter
CC (LST) PiT-1 (see AAW70498). Human PiT-1 is also a sodium-phosphate
CC cotransporter and a human amphotrophic retrovirus receptor
CC (leukaemia virus receptor 1). LSTs such as PiT-1 provide the
CC physiological mechanism for the extrusion of lithium from cells,
CC i.e. it regulates the cell concentration of lithium. Its activity
CC determines the therapeutic effect of lithium. The invention
CC provides a simple molecular biological test for the ability of
CC cells to extrude lithium. The LSTs have significance for
CC determining the responsiveness of humans with mental disorders,
CC including manic depressives, to treatment with lithium salts,
CC probes and primers for PiT-1, PiT-2 (see AAV33502) and BNPI (see
CC AAV33503) can be used in diagnostic tests useful for genetic
CC screenings to predict whether a patient will respond to lithium
CC treatment. The test is also a screen for susceptibility to, and
CC extent of, manic depressive illness, and is suitable for screening
CC newborns.

XX SQ Sequence 3220 BP; 784 A; 710 C; 762 G; 964 T; 0 other;

Query Match 96.5%; Score 3173.6; DB 19; Length 3220;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 3181; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 97 CCGACCCGGGCGGTGCGCGGTCCAGCGCTCCAGCGCTCGATCTCTGCTCCCG 156

Db 24 CCGGGCGGTGCGCGGTCCAGCGCTCCAGCGCTCGATCTCTGCTCCCG 83

Qy 157 CTCGCCCTCCCTTTTCCCTGGATCACTTGGCTCTTCTCTCGCCATGGAATTC 216

Db 84 CTCGCCCTCCCTTTTCCCTGGATCACTTGGCTCTTCTCTCGCCATGGAATTC 143

Qy 217 TGCTCCGTGCTTTTAGCCCTCTCTAGCCAAAGAAACCCAGACAGATGCCATACGC 276

Db 144 TGCTCCGTGCTTTTAGCCCTCTCTAGCCAAAGAAACCCAGACAGATGCCATACGC 203

Qy 277 AGCGTATAGCAGTAACCTCCAGCTCGGTTTCTGTCGCGTAGTTTACAGTATTTTATTTT 336

Db 204 AGCGTATAGCAGTAACCTCCAGCTCGGTTTCTGTCGCGTAGTTTACAGTATTTTATTTT 263

Qy 337 ATATAATATATATTTTATTTATAGCATTTTGTACCTCATATTTCTGTTTACATCTT 396

Db 264 ATATAATATATATTTTATTTATAGCATTTTGTACCTCATATTTCTGTTTACATCTT 323

Qy 397 GAAAGCGCTCAGTAGTCTCTTACTAAACACCACTACTCCAGAGAAATGCAACGGTGA 456

Db 324 GAAAGCGCTCAGTAGTCTCTTACTAAACACCACTACTCCAGAGAAATGCAACGGTGA 383

Qy 457 TTACAGTACTACAGTGTCTACCGCGCTTCTGTCGCTTTTGTGGACTACCTATGATGTC 516

Db 384 TTACAGTACTACAGTGTCTACCGCGCTTCTGTCGCTTTTGTGGACTACCTATGATGTC 443

Qy 517 TCATCTCGGCTTCATTTGTCATTTGTCATTTCTCCGTCGAGCAATGATGATG 576

Db 444 TCATCTCGGCTTCATTTGTCATTTGTCATTTCTCCGTCGAGCAATGATGATG 503

Qy 577 CAAATTTCTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCTGAAGCAAGCCTGCATCC 636

QY 3157 GATTTCTGCGAGTGTGGGATGATGAATGAAGTGAATGTGAACCTTTGGGCAAGTTAAAT 3216
DB 3083 GATTTCTGCGAGTGTGGGATGATGAATGAAGTGAATGTGAACCTTTGGGCAAGTTAAAT 3142
QY 3217 GGGACAGCCCTTCCATGTTTCATTTGTCTACCTCTTAACCTGAATAAAAAAGCCCTACAGTTT 3276
DB 3143 GGGACAGCCCTTCCATGTTTCATTTGTCTACCTCTTAACCTGAATAAAAAAGCCCTACAGTTT 3202
QY 3277 TAGAAAAA 3285
DB 3203 TAGAAAAA 3211

RESULT 8

AAQ83679 ID AAQ83679 standard; cDNA; 3211 BP.

XX AC AAQ83679;
XX
DT 30-OCT-1995 (first entry)
XX
DE GALV receptor-encoding cDNA.
XX
KW GALV; receptor; retrovirus; GLVR-1; ss.
XX
OS Gibbon ape leukemia virus.
XX

Key Location/Qualifiers

FT CDS 371..2410

FT /*tag= a

FT mat_peptide 371..2407

FT /*tag= b

FT repeat_unit 125..135

FT /*tag= c

FT repeat_unit 618..628

FT /*tag= d

XX

PN US5414076-A.

XX

PD 09-MAY-1995.

XX

PD 24-AUG-1989; 89US-0398351.

XX

PF 25-MAR-1991; 91US-0674287.

XX

PR 24-AUG-1989; 89US-0398351.

XX

XX (AMCY) AMERICAN CYANAMID CO.

XX

PI O'Hara BM;

XX

XX WPI; 1995-185136/24.

DR P-PSDB; AAR72969.

XX

XX New DNA encoding gibbon ape leukaemia virus receptor - used to develop

PT prods. to prevent virus infections or to bind and inactivate viruses.

XX

PS Claim 1; Column 13-20; 35pp; English.

XX

XX DNA from HL60 cells (easily infected with GALV) and pSV2gpt were

CC introduced into mouse NIH3T3 cells (which cannot be infected with

CC GALV) to produce primary and secondary transformants. A genomic

CC library constructed from secondary transformants was screened for

CC the presence of clones contg. repetitive DNA using human DNA as a

CC probe. The insert DNA obtd. was used as a probe with a cDNA library

CC prep'd. from HL60 cells to obtain DNA (given in AAQ83679) encoding

CC GALV receptor (AAR72969).

XX

XX Sequence 3211 BP; 781 A; 708 C; 760 G; 962 T; 0 other;

SQ

Query Match

Best Local Similarity 96.4%; Score 3172; DB 16; Length 3211;

Matches 3180; Conservative 0; Mismatches 8; Indels 1; Gaps 1;


```

AAX15171
ID AAX15171 standard; cDNA; 3211 BP.
AC AAX15171;
DT 23-APR-1999 (first entry)
XX cDNA encoding the gibbon ape leukemia virus (GALV) receptor.
DE Human gibbon ape leukaemia virus receptor; GALV; substrate transport;
KW ion transport; immune regulatory activity; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 371..2410
FT FT /*tag= a
FT FT /product= GALV receptor
FT FT /transl_except= (pos: 1253..1255, aa: Lys)
FT FT /transl_except= (pos: 1457..1459, aa: Ser)
XX
PN US874264-A.
XX
PD 23-FEB-1999.
XX
PF 08-MAY-1995; 95US-0436900.
XX
PR 25-MAR-1991; 91US-0674287.
PR 24-AUG-1989; 89US-0398351.
PR 08-MAY-1995; 95US-0436900.
XX
PA (AMCY ) AMERICAN CYANAMID CO.
XX
PI O'Hara BM;
XX
DR WPI; 1999-179981/15.
XX
PT Recombinant gibbon ape leukaemia virus receptor proteins - and
PT plasmids containing nucleotide sequences encoding such proteins
XX
PS Claim 4; Fig 6A-C; 34pp; English.
XX
CC The present sequence encodes a human gibbon ape leukaemia virus
CC (GALV) receptor protein. The cDNA sequence was isolated from HL60
CC cells. The GALV receptor proteins are useful in experimental
CC manipulation of the GALV host, in the analysis of virus/receptor
CC interactions, and in the elucidation of the normal role of the
CC receptor, which is thought to include substrate/ion transport and
CC immune regulatory activity.
XX
SQ Sequence 3211 BP; 781 A; 708 C; 760 G; 962 T; 0 other;

Query Match 96.4%; Score 3172; DB 20; Length 3211;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3180; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 97 CCGACCCGGCCGCTGTCGCCGCTGCTCCAGCCGCTGCCGCTCGATCTCTCTCCCG 156
DB 24 CCGGGCCGTGCTGCTGCCGCTGCTCCAGCCGCTGCCGCTCGATCTCTCTCCCG 83
QY 157 CTCGCCCTCCCTTTTCCCTGGATGAATGCTGCTCTCTCTCTCCGCAATGGAATTC 216
DB 84 CTCGCCCTCCCTTTTCCCTGGATGAATGCTGCTCTCTCTCTCTCCGCAATGGAATTC 143
QY 217 TGCTCCGCTGTTTATAGCCCTCTGAGCCAAAGAAACCCAGACAGATGCCCATACGC 276
DB 144 TGCTCCGCTGTTTATAGCCCTCTGAGCCAAAGAAACCCAGACAGATGCCCATACGC 203
QY 277 AGCGTATAGCAGTAATCTCCCGAGCTCGGTTCTGTCGCGTAGTTTACAGTATTTATTT 336
DB 204 AGCGTATAGCAGTAATCTCCCGAGCTCGGTTCTGTCGCGTAGTTTACAGTATTTATTT 263

```

Db	2424	TAAATTTGTGCTCAATGTTGGGACCATCTTAGGTATCTCTGCTCCCTCGAAGAATGATT	2483
Qy	2557	ACAGTGTAAACAGAAAGACTGACAAGAGCTCTTTTATTTGGAGCCAGAGGAGGAAGTGT	2616
Db	2484	ACAGTGTAAACAGAAAGACTGACAAGAGCTCTTTTATTTGGAG-CAGAGAGGGAAGTGT	2542
Qy	2617	TACTTGTGCTATAAAGCTCTTTTGTGCTAAATATGAATTTGTCTCAAAATTAAGCTGTGTAAA	2676
Db	2543	TACTTGTGCTATAAAGCTCTTTTGTGCTAAATATGAATTTGTCTCAAAATTAAGCTGTGTAAA	2602
Qy	2677	ATAGCCGGGTTCCAGCTGGCTCTGCTGAGGTCCCTTCTCTCTGGGCTGTGAATTCCT	2736
Db	2603	ATAGCCGGGTTCCAGCTGGCTCTGCTGAGGTCCCTTCTCTCTGGGCTGTGAATTCCT	2662
Qy	2737	GTACATATTTCTTACCTTTTGTATCAGGCTTCAATTCACATTAATGTTTTTAAATGTTGCTC	2796
Db	2663	GTACATATTTCTTACCTTTTGTATCAGGCTTCAATTCACATTAATGTTTTTAAATGTTGCTC	2722
Qy	2797	TGAAGATGACTGTGATTTTTTTTTTCTTTTTTAAACCATGAAGAGCCGTTTGACAGAG	2856
Db	2723	TGAAGATGACTGTGATTTTTTTTTTCTTTTTTAAACCATGAAGAGCCGTTTGACAGAG	2782
Qy	2857	CATGCTCTGGTGTGTTGGTTTACCAGCTCTGCCCCTCACATGCACAGGGAATTAACAAC	2916
Db	2783	CATGCTCTGGTGTGTTGGTTTACCAGCTCTGCCCCTCACATGCACAGGGAATTAACAAC	2842
Qy	2917	AAAAATATACTACAACTTCCCTTGTAGTCTCTTATATAAGTAGAGTCTCTGGTACTCTG	2976
Db	2843	AAAAATATACTACAACTTCCCTTGTAGTCTCTTATATAAGTAGAGTCTCTGGTACTCTG	2902
Qy	2977	CCCTCTCTCAGTAGTGGCAGGATCTATTGGGCATATTCGGGAGCTCTTAGAGGGGATGAG	3036
Db	2903	CCCTCTCTCAGTAGTGGCAGGATCTATTGGGCATATTCGGGAGCTCTTAGAGGGGATGAG	2962
Qy	3037	GTTCTTTGAACACAGTGAATAATTAATTAAGTAACTTTTTTTCAGCAGTATTATGACTG	3096
Db	2963	GTTCTTTGAACACAGTGAATAATTAATTAAGTAACTTTTTTTCAGCAGTATTATGACTG	3022
Qy	3097	TTATTGCTAAGAACAAGTAAGAAGAAAAGCCGTTTGGCAACTCTGGTATTCTTTTAA	3156
Db	3023	TTATTGCTAAGAACAAGTAAGAAGAAAAGCCGTTTGGCAACTCTGGTATTCTTTTAA	3082
Qy	3157	GATTCTTGGCAGTGTGGATGGATGAATGAAGTGGAAATGTGAACCTTTGGGCAAGTTAAAT	3216
Db	3083	GATTCTTGGCAGTGTGGATGGATGAATGAAGTGGAAATGTGAACCTTTGGGCAAGTTAAAT	3142
Qy	3217	GGGACAGCTTCCATGTTTCATTTGTCTTACCTCTTAACCTGAATAAAAAGCCTACAGTTTT	3276
Db	3143	GGGACAGCTTCCATGTTTCATTTGTCTTACCTCTTAACCTGAATAAAAAGCCTACAGTTTT	3202
Qy	3277	TAGAAAAAA 3285	
Db	3203	TAGAAAAAA 3211	
RESULT 10			
ABQ91973			
ID	ABQ91973 standard; cDNA; 2916 BP.		
XX			
AC	ABQ91973;		
XX			
DT	30-SEP-2002 (first entry)		
XX			
DE	Human NF-kB activating gene SEQ ID NO 124.		
XX			
KW	Human; NF-kB; nuclear factor kappa B; mouse; antiinflammatory;		
KW	immunomodulator; cytosolic; antineoplastic; osteopathic; neoplastic;		
KW	neuroprotective; anti-HIV; autoimmune disease; cancer; infection;		
KW	bone disease; AIDS; neurodegenerative disease; ischaemic disorder;		
KW	gene; ss.		
XX			
OS	Homo sapiens.		
XX			

OS Homo sapiens.

PN WO200253737-A1.
XX 11-JUL-2002.
XX 25-DEC-2001; 2001WO-JP11389.
XX 28-DEC-2000; 2000JP-0402288.
PR 26-MAR-2001; 2001JP-0088912.
PR 24-AUG-2001; 2001JP-0254018.
XX (ASAH) ASahi KASEI KOGYO KK.
XX Matsuda A, Honda G, Muramatsu S, Nagano Y;
PI WPI: 2002-583617/62.
DR P-PSDB; ABP61485.
XX NF-approximatelykB activating gene and expressed protein, applicable in
XX diagnosis and screening inhibitors or promoters to control excessive
PT activation or inhibition for treating e.g. inflammations, autoimmune
PT diseases and cancer -
XX
PS Claim 4; Page 535-542; 841pp; Japanese.
XX The invention relates to a purified protein (I), comprising one of 90
CC fully defined sequences (ABP61424-ABP61513) or a protein based on any of
CC the sequences but with some amino acids deleted, substituted or added and
CC with a NF-kB (nuclear factor kappa B) activating effect. The protein and
CC encoding gene (ABP61512-ABP62001) are useful in diagnosis and screening
CC inhibitors or promoters to control excessive activation or inhibition
CC and for treating e.g. inflammations, autoimmune diseases, cancers,
CC infections, bone diseases, AIDS, neurodegenerative diseases or ischaemic
CC disorders.
XX
SQ Sequence 2916 BP; 725 A; 606 C; 704 G; 881 T; 0 other;
Query Match 88.5%; Score 2911.2; DB 24; Length 2916;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2913; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 364 TTTTGGATACCTCATATTCGTTTACATCTTTGAAAGCGCTCAGTAGTCTCTTACTA 423
DB 1 TTTTGGATACCTCATATTCGTTTACATCTTTGAAAGCGCTCAGTAGTCTCTTACTA 60
QY 424 AACCACTACTCCAGAGATGGCAACGGTGTATACAGTACTACAGTCTACCGCG 483
DB 61 AACCACTACTCCAGAGATGGCAACGGTGTATACAGTACTACAGTCTACCGCG 120
QY 484 CTTCTGGTCTTTGGTGGACTACTATGGATGCTCATCTGGGCTTCATTATTGCAATTG 543
DB 121 CTTCTGGTCTTTGGTGGACTACTATGGATGCTCATCTGGGCTTCATTATTGCAATTG 180
QY 544 TCTTGGCATCTCCGTTGGGAGCAATGATGTAGCAAAATCTTTTGGTACAGTGTGGGT 603
DB 181 TCTTGGCATCTCCGTTGGGAGCAATGATGTAGCAAAATCTTTTGGTACAGTGTGGGT 240
QY 604 CAGGTAGTACCTCCAGCAAGCTGCAATCTAGTACGATCTTTGAAACAGTGGCT 563
DB 241 CAGGTAGTACCTCCAGCAAGCTGCAATCTAGTACGATCTTTGAAACAGTGGCT 300
QY 664 CTGCTTACTTGGGGGCAAGTGGAGCAACCAATCCGGAAGGGCTTGTATTGACGTGGAGA 723
DB 301 CTGCTTACTTGGGGGCAAGTGGAGCAACCAATCCGGAAGGGCTTGTATTGACGTGGAGA 360
QY 724 TGTACAACTCGACTCAAGGGCTACTGATGCCGGCTCAGTCACTGCTATGTTTGGTCTG 783
DB 361 TGTACAACTCGACTCAAGGGCTCTGATGGCCGGCTCAGTCACTGCTATGTTTGGTCTG 420
QY 784 CTGCTGGCAACTCGTGGCTTCCTTTTGAAGCTCCCTATTTCTGGAACCCATTTGATTG 843
DB 421 CTGCTGGCAACTCGTGGCTTCCTTTTGAAGCTCCCTATTTCTGGAACCCATTTGATTG 480
QY 844 TTGGTGCAACTATTGGTTTCTCCCTCGTGGCAAGGGGAGGAGGTGTCAAGTGGTCTG 903

DB 481 TTGGTGCAACTATTGGTTTCTCCCTCGTGGCAAGGGGAGGAGGTGTCAAGTGGTCTG 540
QY 904 AACTGATATAAAATTTGTGATGCTTGGTTCGCTCCCACTGCTTCTTGGATTAATGCTG 963
DB 541 AACTGATATAAAATTTGTGATGCTTGGTTCGCTCCCACTGCTTCTTGGATTAATGCTG 600
QY 964 GAAATTTATCTCTCTGGTTCGTCATTCATCTCCATAAGGCAGATCCAGTTCCTAATG 1023
DB 601 GAAATTTATCTCTCTGGTTCGTCATTCATCTCCATAAGGCAGATCCAGTTCCTAATG 660
QY 1024 GTTTCGAGCTTTGCCAGTTCATGCTGCAAGTGTGAATAAATCTCTTTTCCATCA 1083
DB 661 GTTTCGAGCTTTGCCAGTTCATGCTGCAAGTGTGAATAAATCTCTTTTCCATCA 720
QY 1084 TGTATATCTGGAGCACCTGCTGGCTTTGACAAATCTCTCTGCTGGGGTACCATCTCA 1143
DB 721 TGTATATCTGGAGCACCTGCTGGCTTTGACAAATCTCTCTGCTGGGGTACCATCTCA 780
QY 1144 TCTCGTGGGATGTCAGTTCCTGTCCTTATGCTGCTGCTTCTTGTATGTCCAGGA 1203
DB 781 TCTCGTGGGATGTCAGTTCCTGTCCTTATGCTGCTGCTTCTTGTATGTCCAGGA 840
QY 1204 TGAAGAGAAAAATTAACAGAGAAAAATAAGTGTAGTCTTCTTGAAGCCCTTAATGAAA 1263
DB 841 TGAAGAGAAAAATTAACAGAGAAAAATAAGTGTAGTCTTCTTGAAGCCCTTAATGAAA 900
QY 1264 AAAGATAGCTTGAAGAGAACCATGAGAAACAAAGTTCCTGCTGGTGTATTTGAAA 1323
DB 901 AAAGATAGCTTGAAGAGAACCATGAGAAACAAAGTTCCTGCTGGTGTATTTGAAA 960
QY 1324 ACAAGCATCTCTGTTTCTGAGGTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1383
DB 961 ACAAGCATCTCTGTTTCTGAGGTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1384 AGAGAACAGTCTCATTTCAAACTTTGGAGATTTGGAGAAAGTCCAGAGAGAGAGAGCTTC 1443
DB 1021 AGAGAACAGTCTCATTTCAAACTTTGGAGATTTGGAGAAAGTCCAGAGAGAGAGAGCTTC 1080
QY 1444 CCAGCTGGCTTGAAGAGAGAAACCAAGTGTAGTACAGTGTGCTGCTGCTGCTGCTGCTGCTG 1503
DB 1081 CCAGCTGGCTTGAAGAGAGAAACCAAGTGTAGTACAGTGTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1504 TGCTTAATGGAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1563
DB 1141 TGCTTAATGGAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1564 GCCACTACAGTATCACACCGTGCATAAAGGATTCGGGCTGTGTAAGAGCTACTCCATA 1623
DB 1201 GCCACTACAGTATCACACCGTGCATAAAGGATTCGGGCTGTGTAAGAGCTACTCCATA 1260
QY 1624 AATTAATCTTGGCAAGGTGGGAGATTCATGGGAGATCCGGTGACAAACCTTTAAGGC 1683
DB 1261 AATTAATCTTGGCAAGGTGGGAGATTCATGGGAGATTCGGTGACAAACCTTTAAGGC 1320
QY 1684 GCAATAATAGCTATATCTTATACCTGGCAATATGTGGATGCTCTGGATTCATTTCC 1743
DB 1321 GCAATAATAGCTATATCTTATACCTGGCAATATGTGGATGCTCTGGATTCATTTCC 1380
QY 1744 GTGCCAAAGAGGTGAAACAGAGGCGGAGAAATGGAGAGCTGATGGCCCTAATGCG 1803
DB 1381 GTGCCAAAGAGGTGAAACAGAGGCGGAGAAATGGAGAGCTGATGGCCCTAATGCG 1440
QY 1804 ACTCCAAGAGCGAATTCGAATGGAGCTTACACAGTTCCTGCAATGCTGTGCTGAC 1863
DB 1441 ACTCCAAGAGCGAATTCGAATGGAGCTTACACAGTTCCTGCAATGCTGTGCTGAC 1500
QY 1864 TTAAGTGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1923
DB 1501 TTAAGTGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
QY 1924 AAGGAGTAAATGGCTCTCTAGAGAAATGGTATGATGATGATGATGATGATGATGATGATG 1983

Db 1561 AAGGAAGTAAGCTCTCTAGGAAGTAAGCTGATGACAGGATTAAGCTGAACTCTCTCTCC 1620
Qy 1984 TCTTCCAGTCTCTGAGATCCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGCTGGCAATG 2043
Db 1621 TCTTCCAGTCTCTGAGATCCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGCTGGCAATG 1680
Qy 2044 ACCTAAGCAATGCCATTTGGGCTCTGCTTTGCTTTATATTTGGTTATGACACAGAGATG 2103
Db 1681 ACCTAAGCAATGCCATTTGGGCTCTGCTTTGCTTTATATTTGGTTATGACACAGAGATG 1740
Qy 2104 TTTCTTCAAAAGTGGCAACACCAATATGCTTCTACTCTATGCTGGTGTGTTGCTATCTGTG 2163
Db 1741 TTTCTTCAAAAGTGGCAACACCAATATGCTTCTACTCTATGCTGGTGTGTTGCTATCTGTG 1800
Qy 2164 TTGGTCTGTGGTGGGGAAGAAGATTTATCAGACCATGGGGAAGGATCTGACACCGA 2223
Db 1801 TTGGTCTGTGGTGGGGAAGAAGATTTATCAGACCATGGGGAAGGATCTGACACCGA 1860
Qy 2224 TCACACCTCTAGTGGCTTCAGTATTAAGCTGGCATCTGCCCTCACTGTGGTGTATGAT 2283
Db 1861 TCACACCTCTAGTGGCTTCAGTATTAAGCTGGCATCTGCCCTCACTGTGGTGTATGAT 1920
Qy 2284 CAATATTTGGCTTCCCATCAGTCAACACATTTGTAAGTGGCTCTGTTGCTGTGTTG 2343
Db 1921 CAATATTTGGCTTCCCATCAGTCAACACATTTGTAAGTGGCTCTGTTGCTGTGTTG 1980
Qy 2344 GCTGCTCCGCTCCGGAAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2403
Db 1981 GCTGCTCCGCTCCGGAAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Qy 2404 GCTTGTGACAGTCCCATTTCTGAGTTATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2463
Db 2041 GCTTGTGACAGTCCCATTTCTGAGTTATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Qy 2464 ATGTCATCTCAGATGTAAGCTGCTGAGATTAAGTGTGCTGCTGCTGCTGCTGCTGCT 2523
Db 2101 ATGTCATCTCAGATGTAAGCTGCTGAGATTAAGTGTGCTGCTGCTGCTGCTGCTGCT 2160
Qy 2524 TCTTAGGTATTCCTGCTCCCTGGAAGATGATTACAGTGTAAACAGAGACTGACAAGAG 2583
Db 2161 TCTTAGGTATTCCTGCTCCCTGGAAGATGATTACAGTGTAAACAGAGACTGACAAGAG 2220
Qy 2584 TCTTTTATTTGGGAGCCAGAGGAGGAGTGTACTGCTGCTATTAACCTGCTTTGCTGCT 2643
Db 2221 TCTTTTATTTGGGAGCCAGAGGAGGAGTGTACTGCTGCTATTAACCTGCTTTGCTGCT 2280
Qy 2644 AAATATGATTTCTCAAAATTTAGCTGCTGTAATAGCCCGGTTCCACTGCTCCTGCT 2703
Db 2281 AAATATGATTTCTCAAAATTTAGCTGCTGTAATAGCCCGGTTCCACTGCTCCTGCT 2340
Qy 2704 GAGGTCCCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2763
Db 2341 GAGGTCCCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Qy 2764 GCTTCAATTCATATGTTTAAATGTTGCTCTCTGAGATGCTGTTGATTTTTTTTCT 2823
Db 2401 GCTTCAATTCATATGTTTAAATGTTGCTCTCTGAGATGCTGTTGATTTTTTTTCT 2460
Qy 2824 TTTTAAACCATGAGAGCGTTTGACAGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2883
Db 2461 TTTTAAACCATGAGAGCGTTTGACAGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Qy 2884 CTCTGCTCCATGACAGGAGTTTAAACAAACAAATATTAACATCAACTTCCCTGCTGTA 2943
Db 2521 CTCTGCTCCATGACAGGAGTTTAAACAAACAAATATTAACATCAACTTCCCTGCTGTA 2580
Qy 2944 GTCTCTTATTAAGTAGAGTCTTGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3003
Db 2581 GTCTCTTATTAAGTAGAGTCTTGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Qy 3004 TTGGCATATTCGGGAGCTTCTTAGAGGATGAGGTTCTTTGAACACAGTGAATAATTTAA 3063
Db 2641 TTGGCATATTCGGGAGCTTCTTAGAGGATGAGGTTCTTTGAACACAGTGAATAATTTAA 2700

Qy 3064 TTAGTAACCTTTTTCGAAGCAGTTTATTGACTGTTATTGCTAAGAGAGTAAAGAA 3123
Db 2701 TTAGTAACCTTTTTCGAAGCAGTTTATTGACTGTTATTGCTAAGAGAGTAAAGAA 2760
Qy 3124 AAAGCCTGTTGGCAATCTTGGTTATTCTTTAAAGATTTCTGCGCAGTGTGGATGA 3183
Db 2761 AAAGCCTGTTGGCAATCTTGGTTATTCTTTAAAGATTTCTGCGCAGTGTGGATGA 2820
Qy 3184 TGAAGTGAATGCTGAACCTTTGGGCAAGTTAAATGGACAGCCTTCCATGTTCTTCT 3243
Db 2821 TGAAGTGAATGCTGAACCTTTGGGCAAGTTAAATGGACAGCCTTCCATGTTCTTCT 2880
Qy 3244 ACCTCTTAACCTGAATAAAAAAGCCCTACAGTTTTTTAG 3279
Db 2881 ACCTCTTAACCTGAATAAAAAAGCCCTACAGTTTTTTAG 2916

RESULT 11
ABI9293
ID ABI9293 standard; cDNA; 3227 BP.
XX
AC ABI9293;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:142.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UUNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
WPI; 2002-034733/04.
DR P-PSDB; ABB57070.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
expression levels of particular genes defined in the specification or
by determining the expression profile of a gene group comprising these
genes -
XX
PS Claim 2; Page 437-442; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
conditions, comprising measuring the expression levels of particular
genes (I) in a test sample or determining the expression profile of a
gene group in the sample comprising genes selected from (I). The method
is useful for examining the ischaemic condition (e.g. compressive
ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
the expression levels of particular genes (ABI9202 to ABI9912, encoding
the protein sequences in ABB57020 to ABB57374) or by determining the
expression profile of a gene group comprising these genes. The
expression levels or expression profiles produced by these genes are
used as an indicator when screening for ischaemic condition-improving
drugs or therapeutics for ischaemic diseases. ABI9913 and ABI9914
represent PCR primers for a mouse ischaemic condition related sequence,
XX which are used in the exemplification of the present invention.
SQ Sequence 3227 BP; 781 A; 740 C; 802 G; 904 T; 0 other;
Query Match 64.8%; Score 2133; DB 24; Length 3227;

Best Local Similarity 84.5%; Pred. No. 0;				Matches 2662; Conservative 0; Mismatches 420; Indels 67; Gaps 21;			
Qy	162	CCCTCCCTTTCCCTGGATGAACCTTGGCGTCTTTCTCTCTCGGCATGGAATTTCTGCTC	221	Db	1141	TTTCTGTCCTTATCTGCTCTGTTCTTTGTATGTCCTCCAGGATGAAGAGAAATTTGAACG	1200
Db	124	CACCTCGTCCGCCAGGATGAACCTTGGCGTCTTTCTCTCTATCCGCCATGGAATTTCTGCTC	183	Qy	1223	AGAAATTAAGTGTAGTCTCTCTGAAAGCCCTTAATGAAAGAAAGAAAGAAATAGCTTGAAGA	1282
Qy	222	CGTGCTTTTAGCCCTCTGAGCAAGAAACCCAGACACAGATGCCCATAC--GCAG	278	Db	1201	AGAAATTAAGTGTAGTCTCTGAAAGTCCCTTAATGAAAGAAAGAAAGAAATAGCTTGAAGA	1260
Db	184	CGTGCTTTTAGCCCTCTGAGCAAGAAACCCAGACACAGATGCCCATAC--GCAG	243	Qy	1283	AGACCATGAAGAAACAAAGTGTCTGTTGGTGTATTTGAAACCAAGCATCTCTGTTTCTGA	1342
Qy	279	CGTATAGCAGTAACCTCCCGAGCTCGGTTCTGTCGGTAGTTTACAGTATTTAAATTTTAT	338	Db	1261	AGACCATGAAGAAACAAAGTGTCTCTGAGAGCGTTGAGCATAGGATCTGTGCTCTGA	1320
Db	244	CGTATAGCAGTAACCTCCCGAGCTCGGTTCTGTCGGTAGTTTACAGTATTTAAATTTAT	303	Qy	1343	GGTAGGCTCTCCACTGTGCTCCCTCCAGGCTGTGTGGTAGAGAGAAACAGTCTCAATCAA	1402
Qy	339	ATAATATATATTTATTTATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTGA	398	Db	1321	GGTAGTGTGTCCACTTGGGCCACTCCGGCTGTGTGGAGGAGAGACGGTGTCAATCAA	1380
Db	304	ATAATATATATTTATTTATAGCATTTT--GATACCTCATTTCCGTTTACACATCTCAA	360	Qy	1403	ACTTGGAGATTTGGAGGAGCTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTTCAAGA	1462
Qy	399	RAAGCGCTCAGTAGTTCTCTTA---CTAAACACACACTACTCCAGAGAA-----444	444	Db	1381	ACTTGGTACCTTGGAGGAGCTCCGGAGGAGAGAGGGCTTCCCA---TGGACCTGAAGA	1437
Db	361	RAAGCGCTCAGTAGTTCTCTTA---CTAAACACACACTACTCCAGAGAA-----444	420	Qy	1463	GGAAACAGCATAGATAGCACCGTGAATGGTGCAGTGCAGTTGCCCTAATGGGAACCTTGT	1522
Qy	445	--TGGCAACGGTATTTACAGTACTACAGCTGCTACCGCGCTTCTGTCTCTTTGGTGA	502	Db	1438	GGAGACAGCATAGACAGCACCATCAATGGTGCAGTGCAGTTGCCCTAATGGGAACCTTGT	1497
Db	421	TGTGGCAACGATTACTAGTACCTAGCTGCTGTACTGCTCCGCTCCACCGAATGA	480	Qy	1523	CCAGTTAGTCAAGCCGTACAGCAACCAATAAATCTCAGTGGCCACTACAGTATCAAC	1582
Qy	503	CTACCTATGGATGCTCATCTCGGCTTCATTTATGCAATTTGCTTGGCAATCTCGTGGG	562	Db	1498	TCAGTTAGTCAAACTGTGAGCAACAGATCAACTCCAGTGGCCACTATCAGTATCACAC	1557
Db	481	CAATCTATGGATGCTCATCTCGGCTTCATTTGCAATTTGCTTGGCAATCTCGTGGG	540	Qy	1583	CGTGATTAAGGATTTCCGGCTGTACAAAGAGCTACTCCATAATTTACATCTTGCCAAGT	1642
Qy	563	AGCAATGATGAGCAAAATCTTTGGTACAGCTGTGGCTCAGGTGTAGTACACCTGAA	622	Db	1558	CGTGACAAAGGATTTCTGGCTTGTACAAAGAGCTGTCTCCATAGTTTACATCTTGCCAAGT	1617
Db	541	AGCAATGATGAGCAAAATCTTTGGTACAGCTGTGGCTCAGGTGTAGTACACCTGAA	600	Qy	1643	GGGAGATTGCATGGAGAGACTCCGGTGAACAAACCCCTTAAGGGCCAAATAATAGTATACTTC	1702
Qy	623	GCAAGCTGATCTAGCTAGCATCTTTGAAACAGTGGCTCTGTCTTACTGGGGGCCAA	682	Db	1618	GGGAGCTGCATGGAGATTCTGGGGAACAGCCCTTGAGAGCCAAACAGCTACACTTC	1677
Db	601	GCAAGCTGATCTAGCTAGCATCTTTGAAACAGTGGCTCTGTCTTACTGGGGGCCAA	660	Qy	1703	CTATACCATGGCAATATCTGGATGCCCTCTGATTTCTTCCGTGCCAAAGAGGTGAACA	1762
Qy	683	AGTAGCGAAACCATCCGGAAGGCTTGATTTGACGTGGAGATGTACAATCGACTCAAG	742	Db	1678	CTACACTATGGCAATATCTGGCATGCCCTCGATTTCAATTCGGTGCACAAAGAGGTGAACA	1737
Db	661	AGTAGCGAAACCATCCGGAAGGCTTGATTTGACGTGGAGATGTACAACGAAACTCAAGA	720	Qy	1763	GAAAGGCAAGAAATGGAGAGAGCTGACATGGCTTAATGCAGACTCCAAAGAGCGAATTCG	1822
Qy	743	GCTACTGATGGCGGCTCAGTCACTGCTATGTTGGTCTGCTGCTGGCAACTGCTGGC	802	Db	1738	AAGGGAGATGAAATGGAGAGCTGACATGGCTTAATGCAGATCAAGAGCGAATTCG	1797
Db	721	TCGTCTATGGCTGGCTCCGTGAGTCTATGTTGGTCTGCTGCTGGCAGCTCGTGGC	780	Qy	1823	AATGACAGATTACACAGTTACTGCAATGCTGTGTCTGACCTTCACTCAGCATCTGAGAT	1882
Qy	803	TTCTGTTTTGAAGCTCCCTATTTCTGGAACCCCATTTGTTGTTGGTCAACTATTGGTTT	862	Db	1798	AATGACAGATTACACAGTTACTGCAATGCTGTGTCTGACCTTCACTCAGCATCTGAGAT	1857
Db	781	TTCTGTTTTGAAGCTCCCTATTTCTGGAACCCCATTTGTTGTTGGTCAACTATTGGTTT	840	Qy	1883	AGACATAGTGTCAAGGAGAGATGGGTCTAGGTGACAGAAAGGAAGTAATGGCTCTCT	1942
Qy	863	CTCCCTCTGGGCAAGGGGAGGAGGTGTCAAGTGGTCTGAATGATGATGATGATGAT	922	Db	1858	GGACATAGTGTGAAGGCTGAGATGGGCTGGGTGACAGAAAGGAAGAGCGTGGCTCTCT	1917
Db	841	CTCCCTCTGGGCAAGGGGAGGAGGTGTCAAGTGGTCTGAATGATGATGATGATGAT	900	Qy	1943	AGAAATAGGTATGACACAGGATTAAGCCTGAGTCTCTCTCTTCCAGTTCTCTGCAGAT	2002
Qy	923	GTCTGGTTCGTGCTCCCACTGCTTCTGGAATATGCTGGAAATTTATCTCTCTGCT	982	Db	1918	TGAAGATGGTATGACACAGGATTAAGCCTGAGTGTCCCTCTCTCTTCCAGTTCTCTGCAGAT	1977
Db	901	GTCTGGTTCGTGCTCCCACTGCTTCTGGAATATGCTGGAAATTTATCTCTCTGCT	960	Qy	2003	CTTTACAGCTGCTTTTGGGTCAATTCGCCCATGGTGGCAATGACGTGAAGCAATGCCATTGG	2062
Qy	983	TGCTGATTCATCTCTCAAGGAGATCCAGTTCCTAATGGTTCGAGCTTTGCGAGT	1042	Db	1978	CTTTACAGCTGCTTTTGGGTCAATTTGCCCATGGTGGCAATGACGTGAGCAATGCCATCG	2037
Db	961	TGCTGATTCATCTCTCAAGGAGATCCAGTTCCTAATGGTTCGAGCTTTGCGAGT	1020	Qy	2063	GGCTCTGGTTCCTTATATTTGGTTCATGACAGAGATGTTTCTTCAAGAGTGGCAAC	2122
Qy	1043	TTTCTATGCTGCAGAGTTGGAATAAACCTCTTTTCCATCATGATATCTGGAGCACCGTT	1102	Db	2038	CCCTCTGGTTCCTTATCTTTGTTAT---AAACAAGAAAGCTCTCAAAAGAGCGCAAC	2094
Db	1021	TTTCTATGCTGCAGAGTTGGAATAAACCTCTTTTCCATCATGATATCTGGAGCACCGTT	1080	Qy	2123	ACCAATATGGCTTCTACTCTATGGTGGTGGTATCTGTGTGGTGGTGGTGGG	2182
Qy	1103	GCTGGGCTTTGACAACTTCTCTGTTGGGTACCATCTCATCTCGGTGGATGTCAGT	1162	Db	2095	ACCCATATGGCTTCTGCTTTATGGTGGTGGTGGTGGTGGTGGTGGG	2154
Db	1081	GCTGGGCTTTGACAACTTCTCTGTTGGGTACCATCTCATCTCGGTGGATGTCAGT	1140	Qy	2183	GAAGAGAGTTATCCAGACCATGGGGAAGGATCTGACACCGATTCACACCTCTTAGTGGCTT	2242
Qy	1163	TTTCTGTCCTCTATCTGCTGTTCTTTGTATGTCCTCCAGGATGAAGAGAAATTTGAACG	1222	Db	2155	GAAGAGAGTTATCCAGACCATGGGGAAGGATCTGACACCGATTCACACCTCTTAGTGGT	2214
Db				Qy	2243	CAGTATTTGAACCTGGCATCTGCTCCCTCAGTGTGGTGTGATTTGATCAAAATATTTGGCTTCCCAT	2302
				Db	2215	CAGTATTTGAACCTGGCTTCCCTTAACTGTGTGTATCGCATCAAAATTTGGGCTTCCCAT	2274

```
QY 2303 CAGTCAACACATTTGTAAGTGGGCTCTCTTGTGTCGTGCTGGCTCCGCTCCGATCAAGAA 2362
DB 2275 CAGCAACACATTTGTAAGTGGGCTCTCTTGTGTCGTGCTGGCTCCGATCAAGAA 2334
QY 2363 GCCTGTTGACTGCGCTCTCTTGTGTAACATTTTATGGCTGGTGTGTCACAGTCCCAT 2422
DB 2335 GCCTGTTGACTGCGCTCTCTTGTGTAACATTTTATGGCTGGTGTGTCACAGTCCCAT 2394
QY 2423 TTCTGGAGTTATCAGTGTGCCATCATGGCAATCTTTCAGATATGTCATCTCAGAAATGTG 2482
DB 2395 CTCTGGGTTATCAGTGGCGCTATCATGGCAGTATTTCAAGTACATCATCTCCAGTGTG 2454
QY 2483 AAGCTGTTGAGATTTAAATTTGTGTCATATGTTGGGACCATCTTAGGTATTCCTGCTCC 2542
DB 2455 ACGCT----GGGGTTGAAAGCTGTGTGTCAGTGTCTGGGACCATTTACACATTCCTG-TTC 2509
QY 2543 CCTGAAGATGATTAAGTGTGTAACAGAGAGACTGACAGAGCTTTTATTTGGGAGCA 2602
DB 2510 CTAGAGAGAGCTACAGTGTGTT-GCTGAAGAGAGGCAAGGGTCT-----TAAAGGAGCG 2563
QY 2603 GAGGAGGAGGAGTGTACTTGTGCTATACTGCTTTTGTCTAAATATGAAATGTCTCAA 2662
DB 2564 TGGGAAGGAGTAAATTACACTATAATTTGCTTTTGTCTAAATATGACTTATCTCAA 2623
QY 2663 ATTAGCTGTGTAATAGCCCGGTTCCACTGGCTCCTGCTGAGGTCCCTTTTCCCTTCTG 2722
DB 2624 ATTAGCTATGTAATAGCCAGGTTTCCATTGATTCCTCAAGGT-CCCTTTTCTCTG 2682
QY 2723 GCCTGTGATTCCTGTACATATTTCTCTACTTTTGTATCAGGCTCAATTCATATGT 2782
DB 2683 GCCTATGAATTCCTGTACATATTTCTCTAC-TTTTGTATCAGGCTCAATTCAGTATGT 2741
QY 2783 TTTAATGTTGCTCTGAAGATGACTGTGATTTTCTTTTCTTTTAAACCATGAAGA 2842
DB 2742 TTTAATGTTGCTGTGA-----GATAACTTAGGTGGTCTTTTAAACAGCAGCA 2793
QY 2843 GCCGTTTGACAGAGATGCTGCGTGTGTTTACCAGGTTCTGCCCTCAGATGCAC 2902
DB 2794 GAGCCATTTGATGGCATGTACTGCTTGTGCGGCTCACCAGCTTCTTCCCAACATGCAC 2853
QY 2903 AGGGATTTAAACAACAAATATACTACAACTTCCCTTGTAGTCTCTATATAAGTAGAG 2962
DB 2854 AGGGATTTAAACAACATGTAACTGAAGCTTCCCTCCTCATAGTCTCTCATAGAAATAGTC 2913
QY 2963 TCTTTGGTACTCTGCCCTCCTGTGCTAGTGTGCGAGATCTATTGGCATATTGCGGAGCTT 3022
DB 2914 AC---GGCACTCTGCTC-CTGTCTACTAGTGTGCGAGTCTGTG--ATGTGTGCAACTT 2967
QY 3023 CTTAGAGGATGAGTCTTTTGAACACAGTGAATTTAAAT---TAGTAACCTTTTTCG 3079
DB 2968 CTTAGAGGCGCGAGAATCTTTG-GCAGAGTGGAAATATAAGTTTGTAGTAACCTCTTTTCG 3026
QY 3080 AAGCAGTTTATGACTGTTATTGCTTAAGNAGAGTAAAGAAAGAAAGGCTGTGGCAAT 3139
DB 3027 AAACAGTTTACAGGAC--ATGTTGCTTAAGAAAGCAGGAGACAAA-----GCCCTGGCGGT 3079
QY 3140 CTTGGTATTCTTTAAGATTTCTGCGAGTGTGGATGATGAATGAAGTGAATGTGAA 3199
DB 3080 TGTGGTATTCTTCTGAGATTTCTGCGAGTGTGGATGGGTGAATGAAGTGAATGTGAA 3139
QY 3200 CTTTGGCAAGTTAAATGGGACAGCTTCCATGTTTCATTGTTCTACCTCTTAACTGAATA 3259
DB 3140 CTTTGGCAAAATTAATGGGACAGCTTCCATGTTTCATCTGCTACCTCTTAACTGAAT- 3198
QY 3260 AAAAGCCTACAGTTTATAGAAAAAAA 3288
DB 3199 AAAAGCCTACAGTTTATAGAAAAAAA 3227
```

RESULT 12

AAQ83680

ID AAQ83680 standard; cDNA; 3260 BP.

```
XX AAQ83680;
XX AC
XX 30-OCT-1995 (first entry)
XX GLVR-1-encoding cDNA.
XX GLVR-1; GALV; receptor; retrovirus; ss.
XX Gibbon ape leukemia virus.
XX Key Location/Qualifiers
XX CDS 443..2488
XX mat_peptide /*tag= a
XX repeat_unit /*tag= b
XX repeat_unit 129..139
XX repeat_unit /*tag= c
XX repeat_unit 620..630
XX repeat_unit /*tag= d
XX US5414076-A.
XX PN
XX 09-MAY-1995.
XX PD
XX 24-AUG-1989; 89US-0398351.
XX PF
XX 25-MAR-1991; 91US-0674287.
XX PR
XX 24-AUG-1989; 89US-0398351.
XX (AMCY ) AMERICAN CYANAMID CO.
XX PA
XX O'Hara BM;
XX PI
XX WPI; 1995-185136/24.
XX DR
XX P-PSDB; AAR72970.
XX
XX New DNA encoding gibbon ape leukaemia virus receptor - used to develop
XX prods. to prevent virus infections or to bind and inactivate viruses
XX
XX Disclosure; Column 23-32; 35pp; English.
XX
XX A mouse thymus library was screened with a GLVR-1-specific probe.
XX cDNA clones were obtained that covered an ORF and flanking
XX sequences. A composite of these sequences is given in AAQ83680.
XX The encoded protein (AAR72970) differed at less than 10% of residues
XX from the human GALV receptor protein (AAR72969).
XX
XX Sequence 3260 BP; 790 A; 747 C; 809 G; 914 T; 0 other;
XX
Query Match 64.8%; Score 2133; DB 16; Length 3260;
Best Local Similarity 84.5%; Pred No. 0;
Matches 2662; Conservative 0; Mismatches 420; Indels 67; Gaps 21;
QY 162 CCCTCCCTTTTCCCTGGATGAACCTTGGCTCTCTTCTTCTCCGCAATGGAATTCGTCTC 221
DB 157 CACTCTGCTCCCTCCAGGATGAACCTTGGCTCTTCTTCTTAATCCGCAATGGAATTCGTCTC 216
QY 222 CGTGTGTTTACCCCTCTCCAGCCAAAGAAAGAAAGCCAGACAGATGCCCATAC---GCAG 278
DB 217 CGTGTGTTTACCCCTCTCCAGAGCCAAAGAAAGAAAGCCAGACAGATGCCCATAC 276
QY 279 CGTATAGCAGTAACCTCCCGCAGCTGGTTCGTGCGGTAGTTTACAGTATTTAATTTTAT 338
DB 277 CGTATAGCAGTAACCTCCCGCAGCTGGTTCGTGCGGTAGTTTACAGTATTTAATTTTAT 336
QY 339 ATAATATATATATTTTATTTATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTTGA 398
DB 337 ATAATATATATATTTTATTTATAGCATTTT---GATACCTCATTCGTTTACACATCTCAA 393
QY 399 AAGCGCTCAGTAGTTCCTTTA---CTAAACACACCACTACTCCAGAGAA-----444
DB 394 AAGCGCTTAGTAAATTTCTTTATTATTATTTAAAGAACCACTACTACAGAGATGAATCTAC 453
```


QY 445 --TGGCAAGGTGATTACAGTACTACAGCTGCTACCGCGCTTCTGGTCCCTTTGGTGA 502
DB 454 TGTGGCAAGGATTACTAGTACCTAGCTGCTGTACTGCTTCGCTCCACCGAAGTATGA 513
QY 503 CTACCTATGATGCTCATCTCGGCTTCATATTGCAATTGCTTGGCACTTCCTCGGTGG 562
DB 514 CAATCTATGATGCTCATCTCGGCTTCATATTGCAATTGCTTGGCACTTCCTCGGTGG 573
QY 563 AGCCAAATGATGAGCAAAATCTTTTGGTATACAGCTGCTGGCTCAGGTGATGACCTGAA 622
DB 574 AGCCAAATGATGAGCAAAATCTTTGCTGATACAGCTGCTGGCTCAGGTGATGACCTGAA 633
QY 623 GAAGCTGCTATCTAGCTAGCATCTTTGAAACAGTGGCTCTGCTTACTGGGGCCAA 682
DB 634 GAAGCTGCTATCTAGCTAGCATCTTTGAAACAGTGGCTCTGCTTACTGGGGCCAA 693
QY 683 AGTGAGCGAAACCATCCGGAAGGCTTGATTGACGTGGAGATGTACAACTCGACTCAAG 742
DB 694 AGTGAGCGAAACCATCCGGAAGGCTTGATTGACGTGGAGATGTACAACTCGACTCAAG 753
QY 743 GCTACTGATGGCGGCTCAGTCACTGCTATGTTGGTCTGCTGCTGGCAACTCTGCTGGC 802
DB 754 TCTGCTCATGGCTGGCTCGCTCAGTCACTGCTATGTTGGTCTGCTGCTGGCAGCTCGTGGC 813
QY 803 TTGCTTTTGAAGCTCCCTATTCTGGAACCCATTGTTGTTGGTGGCAACTATTGGTTT 862
DB 814 TTGCTTTTGAAGCTCCCTATTCTGGAACCCATTGTTGTTGGTGGCAACTATTGGTTT 873
QY 863 CTCCCTCGTGGCAAGGCGCAGGAGGTGTCAAGTGGTCTGAACCTGATATAAAATTTGTAT 922
DB 874 CTCCCTCGTGGCAAGGCGCAGGAGGTGTCAAGTGGTCTGAACCTGATATAAAATTTGTAT 933
QY 923 GTCTTGGTTCGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 982
DB 934 GTCTTGGTTCGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 993
QY 983 TGTGCTATTCATCTCCATATAGGAGATCCAGTTCCTAATGTTGCTGAGCTTCCGAGT 1042
DB 994 TGTGCTATTCATCTCCATATAGGAGATCCAGTTCCTAATGTTGCTGAGCTTCCGAGT 1053
QY 1043 TTTCTATGCTGTCACAGTTGGAATAAACCCTCTTTTCCATCATGTATACCTGAGCACCGTT 1102
DB 1054 TTTTATGCTGTCACAAATCGGAATCAACCTCTTTTCCATCATGTATACCTGAGCACCGTT 1113
QY 1103 GCTGGCTTTGACAAACTTCTCTGCTGGGTACCATCTCATCTCGGTGGGATGTGCACT 1162
DB 1114 GCTGGCTTTGACAAACTTCTCTGCTGGGTACCATCTCATCTCGGTGGGATGTGCACT 1173
QY 1163 TTTCTGTCCTTATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1222
DB 1174 TTTCTGTCCTTATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1233
QY 1223 AGAAATAAAGTGTAGTCTCTGAAAGCCCTTAATGGAAAAAAGAAATAGCTTGAAGA 1282
DB 1234 AGAAGTAAAGTGTAGTCTCTGAAAGTCCCTTAATGGAAAAAAGAAATAGCTTGAAGA 1293
QY 1283 AGACCATGAAGAAACAAAGTTGCTGTTGGTATATGAAACAAAGCATCTCTGTTCTGA 1342
DB 1294 AGACCATGAAGAAACAAAGTTGCTGTTGGTATATGAAACAAAGCATCTCTGTTCTGA 1353
QY 1343 GGTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402
DB 1354 GGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
QY 1403 ACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGGCTTCCAGCGCTGACCTTGAAGA 1462
DB 1414 ACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGGCTTCCAGCGCTGACCTTGAAGA 1470
QY 1463 GGAACACGATATAGTACCACTGATGTTGCTGATGTTGCTGATGTTGCTGATGTTGCTGAT 1522
DB 1471 GGAGACCATAGACAGACCATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1530

QY 1523 CCAGTTTCAAGCCGTGAGCAACCAATAAATCTCCAGTGGCCACTACAGTATCACAC 1582
DB 1531 TCAGTTTCAAGCCGTGAGCAACCAATAAATCTCCAGTGGCCACTACAGTATCACAC 1590
QY 1583 CGTGCATTAAGGATTTCCGGCTGTACAAAGAGCTACTCCATAAATTTACATCTTGGCAAGGT 1642
DB 1591 CGTGCATTAAGGATTTCCGGCTGTACAAAGAGCTACTCCATAAATTTACATCTTGGCAAGGT 1650
QY 1643 GGGAGATTCATGGGAGACTCCGGTGACAAACCCCTTAAGGGCAATATAGCTATACATTC 1702
DB 1651 GGGAGATTCATGGGAGACTCCGGTGACAAACCCCTTAAGGGCAATATAGCTATACATTC 1710
QY 1703 CTATACCATGCAATATGTGGCATGCTCTGATTTCAATTCCTGTCGCAAGAGGTGAACA 1762
DB 1711 CTACACTATGCAATATGTGGCATGCTCTGATTTCAATTCCTGTCGCAAGAGGTGAACA 1770
QY 1763 GAAGGCGAAGAAATGGAGAAAGCTGACATGCGCTTAATGCAGACTCCAAAGAGCGAATTCG 1822
DB 1771 AAAGGAGATGAAATGGAGAGCTGACATGCGCTTAATGCAGATACCAAGAGCGAATTCG 1830
QY 1823 AATGCAGATTTACACCAAGTTACTGCAATGCTGCTGACCTTCACTCAGCATCTGAGAT 1882
DB 1831 AATGCAGATTTACACCAAGTTACTGCAATGCTGCTGACCTTCACTCAGCATCTGAGAT 1890
QY 1883 AGACATGATGCTCAAGGAGAGATGGGTCTAGGTGACAGAAAGAGAAAGTAAATGGCTCTCT 1942
DB 1891 GGACATGATGTTGAAGCTGAGATGGGCTGGGTGACAGAAAGAGAGAGTGGCTCTCT 1950
QY 1943 AGAAGATGGTATGACACAGGATAAGCTGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2002
DB 1951 TGAAGATGGTATGACACAGGATAAGCTGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2010
QY 2003 CCTTACACCTGCTTTGGGTCACTCGCCCATGGTGGCAATGACGTAAAGCAATGCCATGG 2062
DB 2011 CCTTACACCTGCTTTGGGTCACTTTGCCCATGGTGGCAATGACGTAAAGCAATGCCATGG 2070
QY 2063 GCCTCTGCTGCTTTATATTTGGTTATGACACAGAGATGTTCTTCAAAAGTGGCAAC 2122
DB 2071 CCCTCTGCTGCTTTGTTATCTTGTATAT --- AAACAAGAGCCCTCTACANAAGCGGCAAC 2127
QY 2123 ACCAATATGGCTTTACTCTATGTTGGTGTGTTATCTGTTGTTGTTGTTGTTGTTGTTGTTG 2182
DB 2128 ACCAATATGGCTTTACTCTATGTTGGTGTGTTATCTGTTGTTGTTGTTGTTGTTGTTGTTG 2187
QY 2183 AAGAGAGTTATCCAGACCATGGGAGAGAGCTGACCCCAATCAACCCCTCCAGCTTCCAT 2242
DB 2188 AAGAGAGTTATCCAGACCATGGGAGAGAGCTGACCCCAATCAACCCCTCCAGCTTCCAT 2247
QY 2243 CAGTATTCGAACCTGGCATCTGCCCTCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2302
DB 2248 CAGTATTCGAACCTGGCATCTGCCCTCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2307
QY 2303 CAGTACAAACATTTGAAAGTGGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2362
DB 2308 CAGTACAAACATTTGAAAGTGGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2367
QY 2363 GCTGTTGCTGCGGCTCTCTTTCGTAACATTTTATGCGCTGTTTTCACAGATCCCAT 2422
DB 2368 GCTGTTGCTGCGGCTCTCTTTCGTAACATTTTATGCGCTGTTTTCACAGATCCCAT 2427
QY 2423 TTTCTGGAGTTATCAGTGTGCTGCTCATGCGCAATCTTCAGATATGTCATCCTCAGAAATGTG 2482
DB 2428 CTCTGGGTTATCAGTGTGCTGCTCATGCGCAATCTTCAGTATATCAAGTATCATCTCCAGTGTG 2487
QY 2483 AAGCTGTTGAGATTTAAATTTGCTCAATGTTTGGGACCATCTTTAGGTATTTCTGCTCC 2542
DB 2488 AAGCT --- GGGGTTGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2542
QY 2543 CTTGAGAGATGATTAAGTGTGTTAAACAGAGCTGACAGAGCTTTTATTTTGGGAGCA 2602
DB 2543 CTTGAGAGATGATTAAGTGTGTTAAACAGAGCTGACAGAGCTTTTATTTTGGGAGCA 2596
QY 2603 GAGGAGGGAAGTGTACTTGTGCTATAAATTTGCTGCTATAAATGAAATTTGCTCAAA 2662

[illegible]

RESULT 13

[illegible]

25-MAR-1991; 91US-0674287.
24-AUG-1989; 89US-0398351.
08-MAY-1995; 95US-0436900.
(AMCY) AMERICAN CYANAMID CO.
O'Hara BM;
WPI; 1999-179981/15.
P-PSDB; AAW96999.
Recombinant gibbon ape leukaemia virus receptor proteins - and
plasmids containing nucleotide sequences encoding such proteins
Claim 4; Fig 9A-B; 34pp; English.
The present sequence encodes a murine gibbon ape leukaemia virus
(GALV) receptor protein. The cDNA sequence was isolated from HL60
cells. The GALV receptor proteins are useful in experimental
manipulation of the GALV host, in the analysis of virus/receptor
interactions, and in the elucidation of the normal role of the
receptor, which is thought to include substrate/ion transport and
immune regulatory activity.
Sequence 3260 BP; 790 A; 747 C; 810 G; 913 T; 0 other.

Query Match	64.8%	Score 2131.4	DB 20	Length 3260
Best Local Similarity	84.5%	Pred. No. 0		
Matches 2661	Conservative 0	Mismatches 421	Indels 67	Gaps 21
Qy	162	CCCTCCCTTTTCCCTGGATGAACCTGGTCTCTTCTCTCCGCCATGGAAATCTGCTC	221	
Db	157	CACCTGTCGCCCCAGGATGAACCTGGTCTCTTCTCTAAATCCGCCATGGAAATCTGCTC	216	
Qy	222	CGTGCTTTTAGCCCTCTCGAGCCAAAGAAACCCAGACAACAGATGCCATAC---CGAG	278	
Db	217	CGTGCTTTTAGCCCTCCAGAGCCAAAGAAACCCAGACAACAGAGCCACGACGAG	276	
Qy	279	CGTATAGAGTAATCCGCCAGCTCGGTTTCTGTGCCGTAGTTTACAGTATTTTAT	338	
Db	277	CGTATAGAGTAATCCGCCAGCTCGGTTTCTGTGCCGTAGTTTACAGTATTTTAT	336	
Qy	339	ATAATATATATTATTATATAGACATTTTGTGATACCTCATATCTGTTTACACATCTTGA	398	
Db	337	ATAATATATATTATTATATAGACATTTT---GATACCTCATTCGTTTACACATCTCAA	393	
Qy	399	AAGCGCTCAGTAGTCTCTTA----CTAAACAACCACTACTCCAGAGAA-----	444	
Db	394	AAGCGCTTAGTAATCTCTATTATTAAAGAACCACTACACTAGAGAATGGAATCTAC	453	
Qy	445	--TGGCAACGGTAGTACCAGTACTACAGCTGCTACCGCGCTCTCGTCTCTTTGGTGA	502	
Db	454	TGTGCAACGATTACTAGTACCCCTAGCTGTTTACTGCTTCCGCTCCACCGCAAGTATGA	513	
Qy	503	CTACCTATGATGCTCATCTCGGCTTCATTATTGCAATTTGCTTTGGCATTTCCGCTGG	562	
Db	514	CAATCTATGATGCTCATCTCGGCTTCATCATTTGCAATTTGTCTTGGCATTTCTCGTGG	573	
Qy	563	AGCAATGATGTAGCAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTACCCCTGAA	622	
Db	574	AGCAATGATGTAGCAAAATCTTTTGGTACAGCTGTAGGCTCAGGTGTAGTACCCCTGAA	633	
Qy	623	GCAAGCCTGCATCTCTAGCTAGCATCTTTGAAACAAGTGGGCTCTGTCTTACTGGGGCCAA	682	
Db	634	GCAAGCCTGCATCTTCTAGCTAGCATCTTCGAAACTGTGGGCTTCGCTTGTCTGGGGCCAA	693	
Qy	683	AGTGAGGAAACCATCCGGAAGGCTTGATTGACGTGGAGATGTACAACTCGACTCAAGG	742	
Db	694	AGTGAGGAAACCATCCGGAACGGCTTGTATAGATGTGAGCTGTACAAACGAAATCTCAAG	753	
Qy	743	GCTACTGATGGCCGGCTCAGTCAGTGTATGTTTGGTTCTGCTGTGTGGCAACTCGTGGC	802	

Db 754 TCTGCTCATGGCTGGCTCGCTCAGTGTATGTTTGGTCTCTGCTGTGTGCGAGCTCGTGGC 813
Qy 803 TTTCGTTTTGAGCTCCCTATTTCTGAAACCAATGTAATGTTGGTGAACATTAATGGTTT 862
Db 814 TTTCGTTTTGAGCTTCCGATTTCTGGGACCAATGTAATGTTGGTGAACATTAATGGTTT 873
Qy 863 CTCCCTCGTGGCAAAAGGGGAGGAGGTGTCAAGTGGTCTGAACATGATAAAATTTGTGAT 922
Db 874 CTCCCTGTGGCAATGGGAGGAGGAGGTGTCAAGTGGTCTGAACATGATAAAATTTGTGAT 933
Qy 923 GTCTGGTTCGTGTCCTCCACTGCTTTCTGGAATTAATGTCTGGAATTTAATCTTCCTGCT 982
Db 934 GTCTGGTTCGTGCTCCTCGCTGCTTTCTGATTAATGTCTGGAATTTAATCTTCCTGCT 993
Qy 983 TCGTGATTCATCTCTCCATAAGGCAGATCCAGTTCCTAATGTTTCCGAGCTTTGCCAGT 1042
Db 994 TCGTGAGTTTCATCTCTCCGTAAGGCAGATCCGTTTCTAATGCTTACGAGCTTTACCAAT 1053
Qy 1043 TTTCATGCTGACAGTTGGAAATAACCTCTTTCCATCATGTATATCTGGAGCACCGTT 1102
Db 1054 TTTTATGCTGACAAATCGGAATCAACCTCTTTCCATTAATGTAATCTGGAGCACCGTT 1113
Qy 1103 GCTGGCTTTGACAACTTCTCTGTTGGGTACCATCTCATCTCGGTGGGATGTCAGT 1162
Db 1114 GCTGGCTTTGACAACTTCTCTGTTGGGTACCATCTCATCTCGGTGGGATGTCAGT 1173
Qy 1163 TTTCGTGCTGCTATGCTGCTGTTCTTTGATGTCCTGAGGATCAAGAGAAATTTGAACG 1222
Db 1174 TTTCGTGCTGCTATGCTGCTGTTCTTTGATGTCCTGAGGATCAAGAGAAATTTGAACG 1233
Qy 1223 AGAAATAAAGTGTAGTCTTCTGAAAGCCCTTAATGGAATAAAGAAATAGCTTGAAGA 1282
Db 1234 AGAAGTAAAGTGTAGTCTGCTGAAAGTCCCTTAATGGAATAAAGAAATAGCTTGAAGA 1293
Qy 1283 AGACCATGAAGAAACAAAGTTGTCTGTTGGTGAATTTGAAACAAAGCATCTGTTTCTGA 1342
Db 1294 AGACCATGAAGAAACAAAGTGTCTCTGAGACGTTGAGCATAGGAATCTCTGTTCTGA 1353
Qy 1343 GGTAGGCTGACCTGCTGCTCCCTCCAGCTGTGGTGGAGGAGAACTCTCATTTCAA 1402
Db 1354 GGTAGTGTGCTGCTGCTGCTCCCTCCAGCTGTGGTGGAGGAGAACTCTCATTTCAA 1413
Qy 1403 ACTTGAGATTTTGGAGGAAGCTCCAGAGAGAGAGGCTTCCAGCGTGACCTTGAAGA 1462
Db 1414 ACTTGAGTCTGAGAGGCTCCGAGCGAGAGCGGCTTCCA---TGGACCTGAAGA 1470
Qy 1463 GGAACACGATAGATAGACCGTGAATGGTGCAGTGCAGTTCCTTAATGGGAACCTTGT 1522
Db 1471 GGAGACGAGATAGACAGACCATCAATGGTGCAGTGCAGTTCCTTAATGGGAACCTTGT 1530
Qy 1523 CGAGTTTCAGTCAAGCCGTGAGCAACCAATAAATCACTCCAGTGGCCACTACAGTATCACAC 1582
Db 1531 TCAGTTTCAGTCAAACTGTACAGCAACAGATCAACTCCAGTGGCCACTACAGTATCACAC 1590
Qy 1583 CGTGCAATAGGATTCGGGCTGTACAAAGAGTCTCTCAATAATACATCTTCCCAAGGT 1642
Db 1591 CGTGCAAGGATTCGGCTGTACAAAGAGTCTCTCAATAATACATCTTCCCAAGGT 1650
Qy 1643 GGAGATTCATCGGAGACTCCGGTGAACAAACCTTTAAGCGCAATAAATAGCTATATTC 1702
Db 1651 GGAGACTGATGGGAGATTCGGGACAAAGCCCTTTGAGACGCAACAAACAGCTATACATTC 1710
Qy 1703 CTATACATGGCAATATGGCATGCTCTGGATTCATTCCTGGTCCCAAGAGGTCGAACA 1762
Db 1711 CTACATATGGCAATATGGCATGCTCTGGATTCATTCCTGGTCCCAAGAGGTCGAACA 1770
Qy 1763 GAAGGCGAAGAAATGGAGAGTGTACATGGCCCTAATGAGACTCCAAAGAGCGAATTCG 1822
Db 1771 AAAGGAGATGAATGGAGAGCTGACATGGCCCTAATGAGATACCAAGAGCGGATTCG 1830
Qy 1823 AATGGACAGTTACACAGTTACTGCAATGCTGTGTCGACCTTCACTCAGATCTCAGAT 1882
Db 1831 AATGGACAGTTACACAGTTACTGCAATGCTGTGTCGACCTTCACTCAGATCTCAGAT 1890

Qy 1883 AGACATGAGTGTCAAGGCGAGAGATGGGTCTAGGTGACAGAAAGGAAGTAATGGCTCTCT 1942
Db 1891 GGACATGAGTGTGAAGGCTGAGATGGGCTGGGTGACAGAAAGGAAGCAGTGGCTCTCT 1950
Qy 1943 AGAAGATGGTATGACACAGGATAAGCTGAAAGTCTCTCTCTCTCTCTCTCTCTCTCT 2002
Db 1951 TGAAGATGGTATGACACAGGATAAGCTGAAAGTCTCTCTCTCTCTCTCTCTCTCTCT 2010
Qy 2003 CCTTACAGCTGCTTTGGGTCAATTCGCCCATGGTGGCAATGACGTTAAGCAATGCCATTGG 2062
Db 2011 CCTTACAGCTGCTTTGGGTCAATTTGCCCATGGTGGCAATGACGTTAAGCAATGCCATTGG 2070
Qy 2063 GCTCTGCTGCTTTTATATTTGGTTTATGACACAGGAGATGTTCTTCAAAAGTGGCAAC 2122
Db 2071 CCTCTGCTGCTTTTATATTTGGTTTAT---AAACAAGAGCTCTCAAAAGCGGCAAC 2127
Qy 2123 ACCAATATGGCTTCTACTCTATGGTGGTGTGGTATCTGTGTGGTCTGTGGGTTTGGGG 2182
Db 2128 ACCCATATGGCTTCTGCTTTATGGTGGTGTGGCATTTGCAATGGGCTGTGGGTTTGGGG 2187
Qy 2183 AAGAAGATTATCCAGACCATGGGGAAGGATCTGACCGATCACACCTCTAGTGGCTT 2242
Db 2188 AAGAAGATTATCCAGACCATGGGGAAGGACCTGACCCCAATCACACCTCTCAGTGGTT 2247
Qy 2243 CAGTATTGAACCTGGCATCTGCCCTCACCTGTGGTGTATGTCATCAAAATATTGGCTTCCC 2302
Db 2248 CAGTATTGAACCTGGCATCTGCCCTCACCTGTGGTGTATGTCATCAAAATATTGGCTTCCC 2307
Qy 2303 CAGTACAAACATTTGAAAGTGGCTCTGTTGTGTCTGTGGTGGCTCCGGTCCAAGAA 2362
Db 2308 CAGCACAAACATTTGAAAGTGGCTCTGTTGTGTCTGTGGTGGCTCCGATCAAGAA 2367
Qy 2363 GGCTGTTGACCTGGCTCTCTTTGTAACATTTTATGGCTGGTGTGTGTCAGATGCCCAT 2422
Db 2368 GGCTGTTGACCTGGCTCTCTTTGTAACATTTTATGGCTGGTGTGTGTCAGATGCCCAT 2427
Qy 2423 TTCTGGAGTTATCAGTGTGCTGCATCATGGCAATCTTCAAGATATGTCATCTCAGAAATGTG 2482
Db 2428 CTCTGGGTTATCAGTGTGCTGCATCATGGCAGTATTTCAAGTACATCTCTGCCAGTGTG 2487
Qy 2483 AAGCTGTTGAGATTTAAATTTGTCTCAATGTTTGGGACCATCTTATAGGTATTCCTGTCC 2542
Db 2488 ACGCT---GGGTTGAAAGCTGTCTAGTGTCTGGACCATTTGATACATTTCTCTG-TTC 2542
Qy 2543 CCTGAAGATGATTAAGTGTAAACAGAGACTGACAGAGTCTTTTTATTTTGGGAGCCA 2602
Db 2543 CTAGAGAGACGCTCACAGTG-TGGCTGAAGACAGCAAGGCTCT-----TAAAGGAGCGG 2596
Qy 2603 GAGGAGGGAAGTGTACTTTGTCTATACTGCTTTTGTCTAAATATGAATTTGTCTCAAA 2662
Db 2597 TGGGAAGGAAGTGTAAATTTACACTAATATGCTTTTGTGCTAAATATGACTTATCTCAA 2656
Qy 2663 ATTAGCTGTGTAATAATAGCCGGTTCCACTGGCTCTGCTGAGTCCCTTTTCTCTCTG 2722
Db 2657 ATTAGCTGTGTAATAATAGCCAGTTTCCATTGATTTCAATTTCTCTCTG 2715
Qy 2723 GGCTGTGAATTCCTGTACATATTTCTACTTTTGTATCAGGCTTCAATTCATTTATGT 2782
Db 2716 GGCTGTGAATTCCTGTACATATTTCTCTAC-TTTTGTATCAGGCTCAATTCAGTATGT 2774
Qy 2783 TTTAATGTTGCTCTGGAAGATGACTTGTGATTTTTTTTTTTTTTTTTTAAACCATGAAGA 2842
Db 2775 TTTAATGTTGCTGTA-----GATAACTTAGGTGGTCTTTTTTAAACAGCCAGCA 2826
Qy 2843 GCCGTTTGACAGAGATGCTCTGCTGTGTGGTTTTCACAGCTTCTGCCCTCACATGCAC 2902
Db 2827 GAGCCATTTGATGGCATGTACTGCTTTGTGGCTCACAGCTTCTTCCCCAACATGCAC 2886
Qy 2903 AGGGATTTAAACAAACAAATAAATACTCAACTTCTCTGTAGTCTCTTATATAAGTAGAG 2962
Db 2887 AGGGATTTAAACAAACATGTAACAGCTTCCCTCTCTCATAGTCTCTCATAGAAATAGTC 2946

QY 1516 ACCTTGTCCAGTTTCAGTCAAGCGCTCAGCAACCAATAAACTCCAGTGGCCACTACCAGT 1575
 Db |||||
 QY 1231 ACCTT-----CGGTTTCAGCGGCCACACAGGAGCGGTCACTGTGT 1273
 Db |||||
 QY 1576 ATCACACCTGTGATTAAGGATTCGGGCTGTACAAGAGAGTACTCCATAAATTACATCTTTG 1635
 Db |||||
 QY 1274 ACCACACCTGTGACAAAGACTCGGGCTGTACAAGAGTCTGTGTGACAAATCCACATCG 1333
 Db |||||
 QY 1636 CCAAGTGGGAGATTTCATGGGAGTCCGGTGACAA-----ACCTTTAAGCGGCAATA 1689
 Db |||||
 QY 1334 ACAGGGGCCCGGAGGAGAGCCAGCGCCAGGAAAGCAATACCGGTGCTCCGCCGAAACA 1393
 Db |||||
 QY 1690 ATAGCTATATCTTCCTATACCATGGCAATATGTGGCATGCTCTGGATTCATTCGCGTCCA 1749
 Db |||||
 QY 1394 ACAGTTACCTGTCTACCGCAGCCATTTTGGGCTGCGAGTGCAGCCACCTTTTCGAG 1453
 Db |||||
 QY 1750 AAGAAGGTGAACAGAGGCGGAGAAATGGAGAAGCTGACATGGCCATAATG---CAGACT 1806
 Db |||||
 QY 1454 CTGCGGACTCATCGGCCCCAGAGGACAGTGAGAAGCTGGTGGCGGACACCGTGTCTACT 1513
 Db |||||
 QY 1807 CCAAGAGCGAATTCGAATGGACAGTTACACGATTTACTGCAATGCTGTCTGACCTTC 1866
 Db |||||
 QY 1514 CCAAGAGAGGCTGCGCTACGACAGCTACTCGAGCTACTGTAAACGCGTGGCAGAGGCGG 1573
 Db |||||
 QY 1867 ACTCAGCATCTGAGA-----TAGCATGAGTGTCAAGGCAGAGATGGGTCTAGGTG 1917
 Db |||||
 QY 1574 AGATCGAGGGGAGGAGGGGGCTGGAGATGAGCTGGGCTGGAGCTGGCGACCTTG 1633
 Db |||||
 QY 1918 ACAGAAAGGAAGTAAATGGCTCTCTAGAAAGATGGTATGACCAAGGATAAGCTTGAAGTCT 1977
 Db |||||
 QY 1634 ACCAGCGCGAGAGGACCTCGCAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1693
 Db |||||
 QY 1978 CTCTCTCTTCAGTTCCTGCGATCCTTACAGCTGCTTTGGGTGATTCGCGCCATGGTG 2037
 Db |||||
 QY 1694 ACCTCTGTTCATTTCTGCAAGTCTTCAACGCTGTTTTCGGGTCTTTTGTCTACGGCG 1753
 Db |||||
 QY 2038 GAATCAGCTAAGCAATGCAATGGGCTCTGGTCTTTATATTTGTTTATGACACAG 2097
 Db |||||
 QY 1754 GCATGACGTGAGTATGCAATCGGTCCTGCTGGTAGCTTGGCTGATTTACAAACAG 1813
 Db |||||
 QY 2098 GAGATGTTTCTTCAAAAGTGGCAACCAATATGGCTTCTACTTATGTTGGTGTGTTA 2157
 Db |||||
 QY 1814 GCGGGGTAAACGCAAGAGCAGCTACACCGCTCTGGCTGTGTTTATGAGGAGTGGAA 1873
 Db |||||
 QY 2158 TCTGTGTTGCTGCTGGGTTTGGGAGAGAGATTATCAGACCATGGGAGGATCTGA 2217
 Db |||||
 QY 1874 TCTGCAAGGCTCTGGGCTCTGGGAGAGAGATGATCCAGACCATGGGAGGACCTCA 1933
 Db |||||
 QY 2218 CACCGATCACACCTCTAGTGGCTTCAGTATTGAACCTGGCATCTGCCCTCACTGTGGTGA 2277
 Db |||||
 QY 1934 CTCCCATCACCGCTCCAGCGGCTTCAGATCGAGCTGGCTCAGCTTCACAGTGGTGA 1993
 Db |||||
 QY 2278 TTGCATCAATATTTGGCTTCCCATCAGTACACCAATTTGAAGTGGGCTCTGTTGTCT 2337
 Db |||||
 QY 1994 TCGCCTCCAACATCGGGCTTCCAGTCAGCACACCGCACTGTAAAGTGGGCTCGGTGGTGG 2053
 Db |||||
 QY 2338 CTGTTGGCTGGCTCCCGTCCAGAGGCTGTGACTGGGCTCTCTTTGCTACATTTTA 2397
 Db |||||
 QY 2054 CCGTGGGCTGGATCCCGCTCCCGAAGGCTGTGAGCTGGGCGCTCTTTCCGAAACATCTCG 2113
 Db |||||
 QY 2398 TGGCCTGGTTGTGTCAGATGCCCCCATTTCTGGAGTATACAGTGTGCGCATCATGGCAATCT 2457
 Db |||||
 QY 2114 TGGCCTGGTTGTGACGCGTCCCTGTGGCTGGGCTGTTCAGCGCTGTGTCATGGCTCTTC 2173
 Db |||||
 QY 2458 TCAGATATGTCATCTCT 2473
 Db |||||
 QY 2174 TCATGATGGATCTCT 2189
 Db |||||

RESULT 15
 AAV33502
 ID AAV33502 standard; cDNA; 3175 BP.
 XX

AC AAV33502;
 XX 18-JAN-1999 (first entry)
 XX Human sodium-lithium countertransporter hPit-2 cDNA.
 DE Sodium-lithium countertransporter; sodium-phosphate cotransporter;
 KW Pit-2; human; leukaemia virus receptor 2; lithium therapy;
 KW manic depression; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 CDS 244..2202
 FT /*tag= a
 FT /product= leukaemia virus receptor 2
 XX WO9838203-A1.
 XX 03-SEP-1998.
 XX 11-FEB-1998; 98WO-US02875.
 XX 27-FEB-1997; 97US-0039462.
 XX (UYEM-) UNIV EMORY.
 XX Gunn, RB, Timmer RT;
 PI WPI; 1998-520759/44.
 DR P-PSDB; AAW70499.
 XX New isolated lithium-sodium counter-transporter DNA - used to
 PT develop products for evaluating lithium-sodium transport in
 PT erythrocytes, particularly for lithium therapy in manic depression.
 XX Claim 5; Page 42-44; 64pp; English.
 CC This DNA molecule encodes human lithium-sodium countertransporter
 CC (LST) Pit-2 (see AAW70499). Human Pit-2 is also a sodium-phosphate
 CC cotransporter and a human amphotrophic retrovirus receptor
 CC (leukaemia virus receptor 2). LSTs such as Pit-2 provide the
 CC physiological mechanism for the extrusion of lithium from cells,
 CC i.e. it regulates the cell concentration of lithium. Its activity
 CC determines the therapeutic effect of lithium. The invention
 CC provides a simple molecular biological test for the ability of
 CC cells to extrude lithium. The LST has significance for
 CC determining the responsiveness of humans with mental disorders,
 CC including manic depressives, to treatment with lithium salts.
 CC Probes and primers for Pit-2, Pit-1 (see AAV33501) and BNPI (see
 CC AAV33503) can be used in diagnostic tests useful for genetic
 CC screenings to predict whether a patient will respond to lithium
 CC treatment. The test is also a screen for susceptibility to, and
 CC extent of, manic depressive illness, and is suitable for screening
 CC newborns.
 XX Sequence 3175 BP; 739 A; 810 C; 815 G; 811 T; 0 other;
 SQ
 Query Match 18.8%; Score 617.2; DB 19; Length 3175;
 Best Local Similarity 60.3%; Pred. No. 1.3e-166;
 Matches 1203; Conservative 0; Mismatches 718; Indels 75; Gaps 8;
 QY 496 TGGTGGACTACCTATGGATGCTCATCTCCCTGGGCTTCATTATTGTCATTTGTCGCTCTCT 555
 Db |||||
 QY 251 TGGATGAGTATTTGGGATGGTCACTTTGGGTTTCATCATAGCTTTCATCTTGGCTTTT 310
 Db |||||
 QY 556 CCGTGGGAGCCCAATGATGTAGCAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGA 615
 Db |||||
 QY 311 CTGTTGGTGCAACCGATGTGCCAATCTCTTTGGTACAGCCGTGGGCTCTGGTGGTGA 370
 Db |||||
 QY 616 CCCTGAAGCAAGCCTGCATCTCTAGTACATCTTTTGAACAGTGGGCTCTGTCTTACTGG 675
 Db |||||
 QY 371 CCTTGAGGAGGATGATGCTTTTAGCTTCAATATTTGAACCAACCGGCTCCGTTACTAG 430
 Db |||||

QY 676 GGGCCAAAGTGAGCGGAAACCATCCGGAAGGGCTTGATTGACGTGGAGATGTACAACCTGGA 735
Db |||||
QY 431 GCGCCAAAGTAGGAGAAACCATTCGCAAAAGGTATCATTTGACGTGAACCTGTGTACAACGAGA 490
Db |||||
QY 736 CTCAAGGGCTACTGATGGCGGCTCAGTCAGTGTCTATGTTTGGTCTCTGCTGTGGCAAC 795
Db |||||
QY 491 CGGTGGAGACTCTCATGGCTGGGAGATTAGTGCCATGGTGGTTCGCGCTGTGGCAGC 550
Db |||||
QY 796 TCGTGGCTTCGTTTTGAAAGCTCCCTATTTCCTGGAACCATTTGTTGTTGTTGCAACTA 855
Db |||||
QY 551 TGATTGCTTCCTTCCTGAGCTTCCAATCTCAGGAACGCACTGCATTGTGGGTTCTACTA 610
Db |||||
QY 856 TTGGTTTCTCCTCGTGGCAAGGGCAGGAGGGTGTCAAGTGGTCTGAACTGTATAAATA 915
Db |||||
QY 611 TAGGAATTTCTCACTGGTTCGCAATCGGTACCAAAAGGTGTGAGTGGATGGAGCTTGTCAAGA 670
Db |||||
QY 916 TTGTGATGTCCTGGTTCGTCGCCACCTGCTTCTGGAATTTATGCTGGAATTTATCT 975
Db |||||
QY 671 TTGTTGCTTCTGGTTTATATCTCCACTGTTGTCTGGTTTCATGCTGGCCTGCTGTTG 730
Db |||||
QY 976 TCTCGTTTCGTGCAATTCATCTCCATTAAGGCAGATCCAGTTCTTAATGTTTGCAGCTT 1035
Db |||||
QY 731 TACTCATCAGAAATTTTCATCTTAAAGGAAGAGACCTGTTCCTCAATGGCCTCCGGCAC 790
Db |||||
QY 1036 TGCCAGTTTCTATGCTGTCACAGTTGGAAATAAACCCTCTTTTCCATCATGTATCTACCTGGAG 1095
Db |||||
QY 791 TCCAGTATTTCTATGCTGCTACCATAGCAATCAATGTCTTTTCCATCATGTATACACAGGAG 850
Db |||||
QY 1096 CACCGTTGCTGGCTTTGACAACTTCTCTGTGGGTACCATCTCATCTCGTGGGAT 1155
Db |||||
QY 851 CACCAAGTCTGGGCTTG---TCTCCCAATGTGGGCCATAGCCCTCATTTCTTTTGGTG 907
Db |||||
QY 1156 GTGCAAGTTTCTGTGCGCTTATCTGCTGCTTCTTGTATGTCCTCAGGATGAAGAGAAAA 1215
Db |||||
QY 908 TCGCCCTCTGTTCTGCTTTTGTGTGGCTCTTCTGTGTCCGTGGATGGAGAGAAAA 967
Db |||||
QY 1216 TTGAACGAGAAATAAAGTGTAGTCTCTTGAAGGCCCTTAAATGGAATAAAGAAATAGCT 1275
Db |||||
QY 968 TAAACAGGCAAAATACAAAAGAAAGGTCTTTATCACAGTATCTGACGAAAGCCCTCAGTA 1027
Db |||||
QY 1276 TGAAGAGAGACCATGAGAAACAAAGTTGCTGTGTTGGTGATTTGAAACAAAGATCTCTG 1335
Db |||||
QY 1028 AGTTTCAGGAAGCAGAGTCCCCAGTATT-----TAAAGAGCTACCCAGG 1070
Db |||||
QY 1336 TTCTGAGTGTAGGCTCTGCCACTGTGCCCTCCAGGCTGTGTGGAGAGAGAACAGTCT 1395
Db |||||
QY 1071 TGCCAAAGCTATGATGACAGACCATCCCGCTCAGGGAGCAGAGGGAG----- 1122
Db |||||
QY 1396 CATTCAAACTTGGAGATTTGGAGGAGCTCCAGAGAGAGAGGCTTCCAGCGTGGACT 1455
Db |||||
QY 1123 -----ACACTGGGGACCTCGGAAGGCACCTCTGCG-----GGCAGCCACCCCTCGGGCT 1170
Db |||||
QY 1456 TGAAGAGGAACACGACATAGATAGCACCGTGAATGGTGTGAGTGTGAGTTGCCATATGGGA 1515
Db |||||
QY 1171 GCATACGGAAGAGCATGTCTCATGACCCATGGCTCTGTGAATCGCCCATCTCCAACGCG 1230
Db |||||
QY 1516 ACCTTGTCCAGTTTCAAGCCGTCAGCAACCAATAAATCACTCCAGTGGCCACTACCACT 1575
Db |||||
QY 1231 ACCTT-----CGGCTTCGAGGGCCACACAGGAGCGCGTATGTGT 1273
Db |||||
QY 1576 ATCACACCGTGCATTAAGGATTCGGGCTGTGTACAAAGAGCTACTCCATAAATTTACATCTTG 1635
Db |||||
QY 1274 ACCACACCGTGCACAAAGACTCGGGCTCTCAAAAGATCTGTGTGACAAATCCACATCG 1333
Db |||||
QY 1636 CCAAGGTGGGAGATTGCATGGGAGACTCCGGTGAACA-----ACCTTTAAGGCGCAATA 1689
Db |||||
QY 1334 ACAGGGGCCCGGAGGAGAGCCAGGCCAGAAAGCAACTACCGGCTGTCTCCGCCGAACA 1393
Db |||||
QY 1690 ATAGCTATATCTTCTATACCATGGCAATATGTGGCATGCTCTGGAATTCATTCCGTGCCA 1749
Db |||||
QY 1394 ACAGTTACCTGCTACACCGCGCAATTTGTGGGCTGCCAGTGCACGCCACCTTTTCGAG 1453
Db |||||

QY 1750 AAGAAAGTGAAACAGAAAGGGCGAAGAAATGGAGAAGCTCACAATGGCCTAATG---CAGACT 1806
Db |||||
QY 1454 CTGGGACTCATCGGCCCCAGAGGACAGTGAAGCTGGTGGCGACACCCGTGTCTACT 1513
Db |||||
QY 1807 CCAAGAACGAATTCGAATGGACAGATTACACCAAGTTACTGCAATGCTGTCTGACCTTC 1866
Db |||||
QY 1514 CCAAGAACGGGTGCGCTACGACAGCTACTCGAGCTACTGTAAACCGGTGGCAGAGCGG 1573
Db |||||
QY 1867 ACTCAGCATCTGAGA-----TAGACATGAGTGTCAAGGCAGAGATGGGTCTAGGTG 1917
Db |||||
QY 1574 AGATCGAGCGGAGGAGGCGCGTGGAGATGAAGCTGGCGTCGGAGCTGGCCGACCTG 1633
Db |||||
QY 1918 ACAGAAAAGGAAGTAATGGCTCTCTAGAAAGTAATGATATGACAGAGATGAAGCTGAGTCT 1977
Db |||||
QY 1634 ACCAGCCGAGAGAGGCCCTCGCAGAGGAGGAGAGGAGAGACGACCCGAGGTTC 1693
Db |||||
QY 1978 CTCTCCTCTTCCAGTTCTCGTGCAGATCCTTACAGCTGCTTTGGGTCAATTCGCCCATGTTG 2037
Db |||||
QY 1694 ACCTCCTGTTCCATTTCTCGAGGTCTCCACCGCTGTTTCGGGTCTTTTGTCTCAGCGG 1753
Db |||||
QY 2038 GCAATGAGCTAAGCAATGCCATTTGGCCTCTGGTTCCTTATATTTGGTTTATGACACAG 2097
Db |||||
QY 1754 GCAATGAGCTGAGTAATGCCATCGTCCCTGGTAGCCTTGTGGCTGATTTACAACAACAG 1813
Db |||||
QY 2098 GAGATGTTTCTTCAAAAGTGGCAACACCAATATGGCTTCTACTCTATGTTGGTGTGTTA 2157
Db |||||
QY 1814 GCGGGTAAACGAAGAAGCAGCTACACCGTCTGGCTGCTGTTTTATGGAGGAGTTGGA 1873
Db |||||
QY 2158 TCTGTGTTGTTGTTGTTGGGTAAGAGATTATCCAGACCATGGGGAAGGATCTGA 2217
Db |||||
QY 1874 TCTGCACAGGCTCTGGGTCTGGGGGAGAGAGTGAATCCAGACCATGGGGAAGGACCTCA 1933
Db |||||
QY 2218 CACGATCACACCTCTAGTGGCTTCAGTATTGAACTGGCATCTGCCCTCACTGTGGTGA 2277
Db |||||
QY 1934 CTCCATCACGCGCTCCAGCGCTTTCAGATCGAGCTGGCTCAGCCTTCAAGTGGTGA 1993
Db |||||
QY 2278 TTGCATCAAAATATTGGCCTTCCCATCAGTACAACAACATTTAAAGTGGGCTCTGTGTGT 2337
Db |||||
QY 1994 TCGCCTCCAAACATCGGGCTTCCAGTCAGCACCAACGCACTGTAAAGTGGGCTCGGTGG 2053
Db |||||
QY 2338 CTGTTGGCTGGCTCGGTCCAGAAAGGCTGTGACTGGCGTCTCTTTGTAACATTTTA 2397
Db |||||
QY 2054 CCGTGGGCTGAGTCGCTCCCGCAAGGCTGTGGACTGGCGCTCTTTCCGGAACATCTTCG 2113
Db |||||
QY 2398 TGGCTGTTGTTGTCACAGTCCCATTTCTGGAGTTATCAGTGTCTGCCATCATGGCAATCT 2457
Db |||||
QY 2114 TGGCTGTTGTTGAGCCGCTCCCTGTGGCTGGGCTGTTTCAGCGTCTGTCATGGCTCTTC 2173
Db |||||
QY 2458 TCAGATATGTCATCCT 2473
Db |||||
QY 2174 TCATGTATGGGATCCT 2189
Db |||||

Search completed: June 28, 2003, 02:59:11
Job time : 492.613 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:39:45 ; Search time 109.651 Seconds
(without alignments)
9201.581 Million cell updates/sec

Title: US-09-981-353-172
Perfect score: 3290
Sequence: 1 cggcgctctcgtggttc.....agttttgagaaaaaaa 3290

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338391 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2.6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2.6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2.6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2.6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2.6/prodata/1/ina/PCUS COMB.seq.*
- 6: /cgn2.6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3172	96.4	3211	1	US-07-674-287B-1
2	3172	96.4	3211	2	US-08-436-900A-1
3	2133	64.8	3260	1	US-07-674-287B-3
4	2131.4	64.8	3260	2	US-08-436-900A-3
5	617.2	18.8	3175	1	US-08-050-684-1
6	617.2	18.8	3175	1	US-08-582-719-1
7	495.2	15.1	662	4	US-09-328-111-623
8	451.4	13.7	613	4	US-09-328-111-521
9	393	11.9	643	4	US-09-328-111-269
10	73.6	2.2	4403765	4	US-09-103-840A-2
11	73.6	2.2	4411529	4	US-09-103-840A-1
12	62.2	1.9	7218	1	US-08-232-463-14
13	58.6	1.8	1029	4	US-09-134-001C-2591
14	52.2	1.6	7218	1	US-08-232-463-14
15	42.2	1.3	289	4	US-09-007-005-17
16	42.2	1.3	289	4	US-09-244-796-17
17	41.2	1.3	4403765	4	US-09-103-840A-2
18	41.2	1.3	4411529	4	US-09-103-840A-1
19	40.4	1.2	31491	4	US-09-360-186-1
20	40.2	1.2	1280	4	US-09-060-756-4
21	39.4	1.2	16442	3	US-08-781-891-208
22	39.2	1.2	521	4	US-09-488-744A-10
23	39	1.2	1028	4	US-08-118-200-1
24	39	1.2	1028	4	US-08-458-745-1
25	38.8	1.2	3999	4	US-09-351-224B-9
26	38.6	1.2	401	4	US-09-221-298-45
27	37	1.1	1283	4	US-09-318-448-22

C 28	36.8	1.1	1181	2	US-08-892-690-2	Sequence 2, Appli
C 29	36.6	1.1	813	4	US-09-288-143-30	Sequence 30, Appli
C 30	36.2	1.1	5923	4	US-09-064-922-3	Sequence 3, Appli
C 31	35.4	1.1	124	6	5185243-1	Patent No. 5185243
C 32	35.2	1.1	930	2	US-08-394-152A-47	Sequence 47, Appli
C 33	34.8	1.1	292	4	US-09-117-121-29	Sequence 29, Appli
C 34	34.8	1.1	292	4	US-09-344-529-8	Sequence 8, Appli
C 35	34.8	1.1	1051	4	US-09-245-041-10	Sequence 10, Appli
C 36	34.8	1.1	2184	4	US-08-955-918C-1	Sequence 1, Appli
C 37	34.8	1.1	2184	4	US-08-697-766A-1	Sequence 1, Appli
C 38	34.8	1.1	3848	4	US-09-112-096-28	Sequence 28, Appli
C 39	34.8	1.1	4081	4	US-08-999-774A-1	Sequence 1, Appli
C 40	34.8	1.1	5668	4	US-09-112-096-14	Sequence 14, Appli
C 41	34.8	1.1	5668	4	US-09-605-785-777	Sequence 777, App
C 42	34.8	1.1	8438	1	US-07-945-283-1	Sequence 1, Appli
C 43	34.6	1.1	1117	4	US-09-247-373B-33	Sequence 33, Appli
C 44	34.6	1.1	1548	2	US-08-762-106-5	Sequence 5, Appli
C 45	34.6	1.1	1548	4	US-09-320-774-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-07-674-287B-1
; Sequence 1, Application US/07674287B
; Patent No. 5414076
; GENERAL INFORMATION:
; APPLICANT: Bryan Mark O'Hara
; TITLE OF INVENTION: Gibbon Ape Leukemia
; TITLE OF INVENTION: Virus Receptor
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Karen A. Lowney
; ADDRESSES: American Cyanamid Company
; STREET: 1937 West Main Street
; CITY: Stamford
; STATE: CT
; COUNTRY: USA
; ZIP: 06904-0060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII converted from IBM DW4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/674,287B
; FILING DATE: 19910325
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A., Dr.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 31,104-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203 321 2361
; TELEFAX: 203 321 2971
; TELEX: 710 474 4059
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3211 base pairs
; TYPE: NUCLEOTIDE SEQUENCE
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: DNA
US-07-674-287B-1

Query Match 96.4%; Score 3172; DB 1; Length 3211;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3180; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy	97	CCGACCCGGGCGGTGTCGCCGTGCGCTCCAGCCGCTGCCGCTCGAATCTCTCTCGGCATGGAATTC	156
Db			
Db	24	CCGGGCGGTGCGGTGTCGCCGTGCGCTCCAGCCGCTGCCGCTCGAATCTCTCTCGCTCTCCCG	83
Qy	157	CTCCGCGCTCCCTTTTCCCTGGATGAACCTTGGCGTCTTCTCTCTCCGCGCATGGAATTC	216
Db	84	CTCCGCGCTCCCTTTTCCCTGGATGAACCTTGGCGTCTTCTCTCTCCGCGCATGGAATTC	143
Qy	217	TGCTCCGCTGCTTTTTCAGCCCTCTCCAGGCCAAGAAACCCAGACAAACAGATGCCCATACGC	276
Db	144	TGCTCCGCTGCTTTTTCAGCCCTCTCCAGGCCAAGAAACCCAGACAAACAGATGCCCATACGC	203
Qy	277	AGCGTATAGCAGTAACCTCCCGAGCTCCGTTCTGTGCGGTAGTTTACAGTATTTAATTTT	336
Db	204	AGCGTATAGCAGTAACCTCCCGAGCTCCGTTCTGTGCGGTAGTTTACAGTATTTAATTTT	263
Qy	337	ATATAATATATATTTATTTATATAGCATTTTTGTATACCTCATATTTCTGTTTTACACATCTT	396
Db	264	ATATAATATATATTTATTTATATAGCATTTTTGTATACCTCATATTTCTGTTTTACACATCTT	323
Qy	397	GAAGGGCGCTCAGTAGTTTCTCTTAATAAACCAACCACTACTCCAGAGAAATGGCAACGGTGA	456
Db	324	GAAGGGCGCTCAGTAGTTTCTCTTAATAAACCAACCACTACTCCAGAGAAATGGCAACGGTGA	393
Qy	457	TTACCAAGTACTACAGCTGCTACCCCGCTCTCGTCTCTTTGGTGGACTACTATGAGATGC	516
Db	384	TTACCAAGTACTACAGCTGCTACCCCGCTCTCGTCTCTTTGGTGGACTACTATGAGATGC	443
Qy	517	TCATCTCGGGCTCATTTATTTGCATTTCTCTGGCATTTCTCCGTGGGAGCAATGATGTAG	576
Db	444	TCATCTCGGGCTCATTTATTTGCATTTGTCTTGGCATTTCTCCGTGGGAGCAATGATGTAG	503
Qy	577	CAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCTCGAAGCAAGCCCTGCATCC	636
Db	504	CAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCTCGAAGCAAGCCCTGCATCC	563
Qy	637	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTCTGGGGCCCAAGTAGCGAAACCA	696
Db	564	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTCTGGGGCCCAAGTAGCGAAACCA	623
Qy	697	TCCGGAAGGGCTTGATTTGACGTGAGATGTACAACTCGACTCAAGGCTACTGATGGCCG	756
Db	624	TCCGGAAGGGCTTGATTTGACGTGAGATGTACAACTCGACTCAAGGCTACTGATGGCCG	683
Qy	757	GCTCAGTCAGTGCTATGTTTGGTTCTGCTGTGTGGCAACTCGTGGCTTCGTTTTGAAGC	816
Db	684	GCTCAGTCAGTGCTATGTTTGGTTCTGCTGTGTGGCAACTCGTGGCTTCGTTTTGAAGC	743
Qy	817	TCCCTATTTCTGGAAACCCATTTGATTTGGTGTCAACTATTTGGTTTCTCCCTCGTGGCAA	876
Db	744	TCCCTATTTCTGGAAACCCATTTGATTTGGTGTCAACTATTTGGTTTCTCCCTCGTGGCAA	803
Qy	877	AGGGCAGGAGGGTGTCAAGTGGTCTGAACGTATAAATAATTTGTCATCTCTTGGTTCTGT	936
Db	804	AGGGCAGGAGGGTGTCAAGTGGTCTGAACGTATAAATAATTTGTCATCTCTTGGTTCTGT	863
Qy	937	CCCCATGCTTTCTGGAAATATGCTGGAATTTTATTTCTTCTGGTTTCGTGCATTCATCC	996
Db	864	CCCCATGCTTTCTGGAAATATGCTGGAATTTTATTTCTTCTGGTTTCGTGCATTCATCC	923
Qy	997	TCCATAAGGCAGATCCAGTTCTTAATGGTTTGCAGCTTTTGCAGTTTTCTATGCGCTGCA	1056
Db	924	TCCATAAGGCAGATCCAGTTCTTAATGGTTTGCAGCTTTTGCAGTTTTCTATGCGCTGCA	983
Qy	1057	CAGTTGGAATAAACCCTCTTTTCCATCATGTATCTCGGAGACCGTCTCGGGCTTTTGACA	1116
Db	984	CAGTTGGAATAAACCCTCTTTTCCATCATGTATCTCGGAGACCGTCTCGGGCTTTTGACA	1043
Qy	1117	AACTTCTCTCTGGGGTACCAATCTCATCTCGGTGGGATGTGCAAGTTTTCTGTGCCCTTA	1176
Db	1044	AACTTCTCTCTGGGGTACCAATCTCATCTCGGTGGGATGTGCAAGTTTTCTGTGCCCTTA	1103

QY	1177	TCGTCCTGGTCTCTTTGTATGTCCACGATGAAGAGAAAAATTTGAACGAGAAATTAAGTGCTA	1233
DB	1104	TCGTCCTGGTCTCTTTGTATGTCCACGATGAAGAGAAAAATTTGAACGAGAAATTAAGTGCTA	1163
QY	1237	GTCCTCTGAAAGCCCTTAATGSAAGAAAGAAATAGCTTTGAAAGAGAACCATGTAAGAAA	1296
DB	1164	GTCCTCTGAAAGCCCTTAATGSAAGAAAGAAATAGCTTTGAAAGAGAACCATGTAAGAAA	1223
QY	1297	CAAAAGTTGTCTGTTGGTGATATTTGAAAACAAGCATCCTGTTTCTGAGGTAAGGCCCTGCCA	1356
DB	1224	CAAAAGTTGTCTGTTGGTGATATTTGAAAACAACCATCCTGTTTCTGAGGTAAGGCCCTGCCA	1283
QY	1357	CTGTGCCCTCCAGGCTGTGGTGAGAGAGAAACAGTCTCATTTCAAACCTTTGGAGATTGG	1416
DB	1284	CTGTGCCCTCCAGGCTGTGGTGAGAGAGAAACAGTCTCATTTCAAACCTTTGGAGATTGG	1343
QY	1417	AGGAAGCTCCAGAGAGAGAGGCTTCCACGCTGGACTTTGAAAGAGGAACACAGCATAG	1476
DB	1344	AGGAAGCTCCAGAGAGAGAGGCTTCCACGCTGGACTTTGAAAGAGGAACACAGCATAG	1403
QY	1477	ATAGCACCGTGAATGGTGCAGTGTGCTTAATGGAAACCTTGTCCAGTTTCAGTCAAG	1536
DB	1404	ATAGCACCGTGAATGGTGCAGTGTGCTTAATGGAAACCTTGTCCAGTTTCAGTCAAG	1463
QY	1537	CCGTGACGAACCAAATAAACTCCAGTGGCCACTACCAAGTATACACCGTGCAATAAGGATT	1596
DB	1464	CCGTGACGAACCAAATAAACTCCAGTGGCCACTCCCAAGTATACACCGTGCAATAAGGATT	1523
QY	1597	CCGGCTGTACAAAGAGCTACTCCATAAATTTACATCTTGCCAGGTGGGAGATTGCATGG	1656
DB	1524	CCGGCTGTGTACAAAGAGCTACTCCATAAATTTACATCTTGCCAGGTGGGAGATTGCATGG	1583
QY	1657	GAGACTCCGGTGACAAACCCCTTAAGGGCCAAATAATAGCTATACTTCTCTATACCATGGCAA	1716
DB	1584	GAGACTCCGGTGACAAACCCCTTAAGGGCCAAATAATAGCTATACTTCTCTATACCATGGCAA	1643
QY	1717	TATGTGCATGCCCTCTGGATTCTATTCGGTGCCAAAGAGGTGAACAGAGGGCGAAGAAA	1776
DB	1644	TATGTGCATGCCCTCTGGATTCTATTCGGTGCCAAAGAGGTGAACAGAGGGCGAAGAAA	1703
QY	1777	TGGAGAAGCTGACATGGCCTTAATGCAGACTCCAGAGAGCGAATTCGAAATGGACAGTTACA	1836
DB	1704	TGGAGAAGCTGACATGGCCTTAATGCAGACTCCAGAGAGCGAATTCGAAATGGACAGTTACA	1763
QY	1837	CCAGTTACTGCAATGCTGTGCTGACCTTCACCTCAGCATCTGAGATAGACATGAGTGTCA	1896
DB	1764	CCAGTTACTGCAATGCTGTGCTGACCTTCACCTCAGCATCTGAGATAGACATGAGTGTCA	1823
QY	1897	AGGCAGAGATGGGTCTAGGTGACAGAAAAGGAATATGGCTCTCTAGAAGAAATGGTATG	1956
DB	1824	AGGCAGAGATGGGTCTAGGTGACAGAAAAGGAATATGGCTCTCTAGAAGAAATGGTATG	1883
QY	1957	ACCAAGATAGCCTGGAAGTCTCTCTCTCTTCAGTTCTCTGAGATCCTTTACAGCCTGCT	2016
DB	1884	ACCAAGATAGCCTGGAAGTCTCTCTCTCTTCAGTTCTCTGAGATCCTTTACAGCCTGCT	1943
QY	2017	TTGGGTCAATTCGCCCATGGTGGCAATCACGTAAGCAATGCCATTTGGSCCTCTGGTTGCTT	2076
DB	1944	TTGGGTCAATTCGCCCATGGTGGCAATCACGTAAGCAATGCCATTTGGSCCTCTGGTTGCTT	2003
QY	2077	TATATTTGGTTTATGACACAGGAGATGTTTCTTCAAAGTGGCAACCAACCAATATGGCTTC	2136
DB	2004	TATATTTGGTTTATGACACAGGAGATGTTTCTTCAAAGTGGCAACCAACCAATATGGCTTC	2063
QY	2137	TACTCTATGGTGGTGGTATCTGTGTTGGTCTGTGGGTTTGGGGAAGAGAGTTATCC	2196
DB	2064	TACTCTATGGTGGTGGTATCTGTGTTGGTCTGTGGGTTTGGGGAAGAGAGTTATCC	2123
QY	2197	AGACCATGGGAAGGATCTGACACCGCATCACCCCTCTAGTGGCTTCAGTATTTGAACCTGG	2256
DB	2124	AGACCATGGGAAGGATCTGACACCGCATCACCCCTCTAGTGGCTTCAGTATTTGAACCTGG	2183
QY	2257	CATCTGCCCTCAGTGTGGTGAATGTCATCAAAATATTTGGCCTTCCCATCAGTACAAACATTT	2316

Db	2184		CATCTGGCCCTCACCTGTGGTGAATGCATCAAAATATTGGCCCTCCCATCAGTACAACACATT	2243
Qy	2317	GTTAAGTGGGCTCTGTGTGTGTCTGTGGCTGGCTCCGGTCCAAGAGCGCTGTTGACGTGGC	2376	
Db	2244	GTTAAGTGGGCTCTGTGTGTGTCTGTGGCTGGCTCCGGTCCAAGAGCGCTGTTGACGTGGC	2303	
Qy	2377	GTCTCTTTTGTGAACAATTTTTTAAGCCCTGGTTGTGCACAGTCCCCAATTTCTGGAGTTATCA	2436	
Db	2304	GTCTCTTTTCGTAAACATTTTTATGGCCCTGGTTGTGCACAGTCCCCAATTTCTGGAGTTATCA	2363	
Qy	2437	GTGCTGCCATCATGGCAATCTTCAGATATGTCATCTCCTCAGNAATGTAAGCTGTTTGAGAT	2496	
Db	2364	GTGCTGCCATCATGGCAATCTTCAGATATGTCATCTCCTCAGNAATGTAAGCTGTTTGAGAT	2423	
Qy	2497	TAAAAATTTGTCTCAATGTTTGGGACCAATCTTAGGTATTCTCTGCTCCCTCGAAGAATGATT	2556	
Db	2424	TAAAAATTTGTCTCAATGTTTGGGACCAATCTTAGGTATTCTCTGCTCCCTCGAAGAATGATT	2483	
Qy	2557	ACAGTGTATAACAGAAGACTGACAAGAGTCTTTTTTATTTGGGAGCCAGAGAGGGAAGTGT	2616	
Db	2484	ACAGTGTATAACAGAAGACTGACAAGAGTCTTTTTTATTTGGGAG-CAGAGAGGGAAGTGT	2542	
Qy	2617	TACTTGTGCTATAACTGCTTTTGTGCTTAATATGAATGTCTCAAAATTAGCTGTGTAAA	2676	
Db	2543	TACTTGTGCTATAACTGCTTTTGTGCTTAATATGAATGTCTCAAAATTAGCTGTGTAAA	2602	
Qy	2677	ATAGCCCGGGTTCACATCGGGCTCCTCTGAGGTGCCCTTCTCTCTGGGCTGTGAAATTCCT	2736	
Db	2603	ATAGCCCGGGTTCACATCGGGCTCCTCTGAGGTGCCCTTCTCTCTGGGCTGTGAAATTCCT	2662	
Qy	2737	GTACATATTTCTCTACTTTTTTGTATCAGGCTTCAATTTCCATTAATGTTTTAATGTGTCTC	2796	
Db	2663	GTACATATTTCTCTACTTTTTTGTATCAGGCTTCAATTTCCATTAATGTTTTAATGTGTCTC	2722	
Qy	2797	TGAAGATGACTGTGATTTTTTTTCTTTTTTCTTTTTTAAACCATGAAGCCGTTTGACAGAG	2856	
Db	2723	TGAAGATGACTGTGATTTTTTTTCTTTTTTCTTTTTTAAACCATGAAGCCGTTTGACAGAG	2782	
Qy	2857	CATGCTCTGGGTTCTGGTTTTCCACGCTTCTGCGCTCACATGCACAGGGAATTTAAACAAC	2916	
Db	2783	CATGCTCTGGGTTCTGGTTTTCCACGCTTCTGCGCTCACATGCACAGGGAATTTAAACAAC	2842	
Qy	2917	AAAAATATACTACAACTTCCTTTGTAGTCTCTTATATAAGTAGAGAGCTCTGGTACTCTG	2976	
Db	2843	AAAAATATACTACAACTTCCTTTGTAGTCTCTTATATAAGTAGAGAGCTCTGGTACTCTG	2902	
Qy	2977	CCCTCCTGTCAGTAGTGGCAGGATCTATTGGCATATTTGGGAGCTCTTAGAGGGGATGAG	3036	
Db	2903	CCCTCCTGTCAGTAGTGGCAGGATCTATTGGCATATTTGGGAGCTCTTAGAGGGGATGAG	2962	
Qy	3037	GTTCTTTTGAACACAGTGAATAATTTAAATTAGTAACTTTTTTTGAAGCAGTTTTATTGACTG	3096	
Db	2963	GTTCTTTTGAACACAGTGAATAATTTAAATTAGTAACTTTTTTTGAAGCAGTTTTATTGACTG	3022	
Qy	3097	TTATTGCTAAGAGAAGTAAGAAAAGAAAAAGCCCTGTGGCAATCTTGGTTATTCTTTTAA	3156	
Db	3023	TTATTGCTAAGAGAAGTAAGAAAAGAAAAAGCCCTGTGGCAATCTTGGTTATTCTTTTAA	3082	
Qy	3157	GATTTCTGGCAGTGGGATGGATGAATGAAGTGAATGTGAACCTTTGGGCAAGTTAAAT	3216	
Db	3083	GATTTCTGGCAGTGGGATGGATGAATGAAGTGAATGTGAACCTTTGGGCAAGTTAAAT	3142	
Qy	3217	GGGACAGCCTTCCAATGTTTCAATTTGCTACCTCTTAACTGAAATAAAAAGCCTACAGTTTT	3276	
Db	3143	GGGACAGCCTTCCAATGTTTCAATTTGCTACCTCTTAACTGAAATAAAAAGCCTACAGTTTT	3202	
Qy	3277	TAGAAAAA 3285		
Db	3203	TAGAAAAA 3211		

RESULT 2

```

US-08-436-900A-1
; Sequence 1, Application US/08436900A
; Patent No. 5874264
; GENERAL INFORMATION:
; APPLICANT: O'Hara, Bryan M.
; TITLE OF INVENTION: Gibbon Ape Leukemia Virus Receptors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,900A
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,104-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-683-2158
; TELEFAX: 201-683-4117
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3211 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-436-900A-1

```

Query Match	96.4%	Score 3172;	DB 2;	Length 3211;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 3180; Conservative	0;	Mismatches	8;	Indels 1; Gaps 1

QY	97	CCGACCGGGCGGTGTGCGCCGTGCGCTCCAGCGGTGCGCGCTCGATCTCCTCGTCTCCGG	156
Db	24	CCGGCGGTGCGGTGCGCGGTGCGCGCTGCGCTGCGCGCTCGATCTCCTCGTCTCCGG	83
QY	157	CTCGCGCCTCCCTTTTCCCTCGATGAACCTGGGTCCTTTCTCTTCTCGCCCATGGAATTC	216
Db	84	CTCGCGCCTCCCTTTTCCCTCGATGAACCTGGGTCCTTTCTCTTCTCGCCCATGGAATTC	143
QY	217	TGCTCCGTCGTCCTTAGCCCTCCTGAGCCAAAGAAACCCAGACACAGATGCCATACGC	276
Db	144	TGCTCCGTCGTCCTTAGCCCTCCTGAGCCAAAGAAACCCAGACACAGATGCCATACGC	203
QY	277	AGCGTATAGCAGTAACCTCCCAGCTCGGTTTCTGTGCGGTAGTTTACAGTATTTAAATTTT	336
Db	204	AGCGTATAGCAGTAACCTCCCAGCTCGGTTTCTGTGCGGTAGTTTACAGTATTTAAATTTT	263
QY	337	ATATAATATATATTATTATTATATAGCATTTTGTGATACCTCATATTTCTGTTTACACATCTT	396
Db	264	ATATAATATATATTATTATTATAGCATTTTGTGATACCTCATATTTCTGTTTACACATCTT	323
QY	397	GAAGGGCGCTCAGTAGTCTCTTACTAAACAACCACTACTCCAGANAATGSCAAACGGTGA	456
Db	324	GAAGGGCGCTCAGTAGTCTCTTACTAAACAACCACTACTCCAGANAATGSCAAACGGTGA	383
QY	457	TTACACGTACTACAGCTGCTACCGCGGCTCTGGTTCCTTTGGTGGACTACCTATGATGC	516
Db	384	TTACACGTACTACAGCTGCTACCGCGGCTCTGGTTCCTTTGGTGGACTACCTATGATGC	443
QY	517	TCATCTCTGGGCTTCATTATTGCATTTTGTCTTTGGCATTTCTCCGTGGGAGCCAATGATGAG	576

Db 444 TCATCTGGGCTTCAATTAATGTCATTTGTCTTGGCAATTCCTCGTGGAGCCAAATGATGTAG 503
Qy 577 CAAATCTCTTTGCTGACAGCTGTGGGCTCAGGTCTAGTGACCCCTGAAGCAAGCCTGCATCC 636
Db 504 CAAATCTCTTTGCTGACAGCTGTGGGCTCAGGTCTAGTGACCCCTGAAGCAAGCCTGCATCC 563
Qy 637 TAGCTAGCACTTTTGAACAGTGGGCTCTGTCTTACTTGGGGGCAAAAGTGAGCGAACA 696
Db 564 TAGCTAGCACTTTTGAACAGTGGGCTCTGTCTTACTTGGGGGCAAAAGTGAGCGAACA 623
Qy 697 TCCGGAAGGCTTGAATGACGTGGAGATGTAACAATCGACTCAAGGGCTACTGATGGCCG 756
Db 624 TCCGGAAGGCTTGAATGACGTGGAGATGTAACAATCGACTCAAGGGCTACTGATGGCCG 683
Qy 757 GCTCAGTCACTATGTTTGGTCTGCTGTGTGGCACTCGTGGCTTCGTTTTTGAAGC 816
Db 684 GCTCAGTCACTATGTTTGGTCTGCTGTGTGGCACTCGTGGCTTCGTTTTTGAAGC 743
Qy 817 TCCCTATTTCTGGAACCCATTTGATTTGTTGGTGCACATATTGGTTCCTCCTCGTGGCAA 876
Db 744 TCCCTATTTCTGGAACCCATTTGATTTGTTGGTGCACATATTGGTTCCTCCTCGTGGCAA 803
Qy 877 AGGGCAGGAGGTGCAAGTGTCTGAACCTGATAAATAATTTGATGTCTGTGTTCTGTGT 936
Db 804 AGGGCAGGAGGTGCAAGTGTCTGAACCTGATAAATAATTTGATGTCTGTGTTCTGTGT 863
Qy 937 CCCCACTGCTTCTGGAATTAATCTCTGGAATTTTATCTCTGCTGCTGCTGCTGCTATTC 996
Db 864 CCCCACTGCTTCTGGAATTAATCTCTGGAATTTTATCTCTGCTGCTGCTGCTGCTATTC 923
Qy 997 TCCATAAGGAGAGATCCAGTTTCTTAATGGTTTGGAGCTTTGGCAGTTTTCATGCTGCA 1056
Db 924 TCCATAAGGAGAGATCCAGTTTCTTAATGGTTTGGAGCTTTGGCAGTTTTCATGCTGCA 983
Qy 1057 CAGTTGGAATAAACCTCTTTTCCATCATGATATATCTGGAGCACCGTGTGCTGGCTTTGACA 1116
Db 984 CAGTTGGAATAAACCTCTTTTCCATCATGATATATCTGGAGCACCGTGTGCTGGCTTTGACA 1043
Qy 1117 AACTCTCTGCTGGGTACCATCTCATCTCGGTGGAGTGTGAGTTCCTGCTGCTGCTTCA 1176
Db 1044 AACTCTCTGCTGGGTACCATCTCATCTCGGTGGAGTGTGAGTTCCTGCTGCTGCTTCA 1103
Qy 1177 TCGTCTGCTGTTTGTATGTCAGGATGAAGAGAAAAATTTGAACGAGAAAAATAAGTGA 1236
Db 1104 TCGTCTGCTGTTTGTATGTCAGGATGAAGAGAAAAATTTGAACGAGAAAAATAAGTGA 1163
Qy 1237 GTCCTTCTGAAAGCCCTTAATGGAAAAAAGAAATAGCTTTGAAAGAACCATGAAAGAA 1296
Db 1164 GTCCTTCTGAAAGCCCTTAATGGAAAAAAGAAATAGCTTTGAAAGAACCATGAAAGAA 1223
Qy 1297 CAAAGTTGCTGTTGGTGATATTGAAACCAAGCATCTGTTTCTGAGGTAGGGCCCTGCCA 1356
Db 1224 CAAAGTTGCTGTTGGTGATATTGAAACCAAGCATCTGTTTCTGAGGTAGGGCCCTGCCA 1283
Qy 1357 CTGTGCCCTTCCAGGCTGTGGTGAGAGAGAACAGTCTCATTTCAAACCTTGGAGATTTGG 1416
Db 1284 CTGTGCCCTTCCAGGCTGTGGTGAGAGAGAACAGTCTCATTTCAAACCTTGGAGATTTGG 1343
Qy 1417 AGGAAGTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGAAACCCAGCATAG 1476
Db 1344 AGGAAGTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGAAACCCAGCATAG 1403
Qy 1477 ATAGCACCGTGAATGGTGAGTGCAGTTGCTTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1536
Db 1404 ATAGCACCGTGAATGGTGAGTGCAGTTGCTTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1463
Qy 1537 CCCTCAGCAACCAATAAATCCAGTGGCCACTACCAAGTATCAACCGTGCATAAGGATT 1596
Db 1464 CCCTCAGCAACCAATAAATCCAGTGGCCACTACCAAGTATCAACCGTGCATAAGGATT 1523
Qy 1597 CCGGCTGTACAAAGAGCTACTCCATAAATTAATCTTGTGCAAGGGTGGAGATTGATGG 1656
Db 1524 CCGGCTGTACAAAGAGCTACTCCATAAATTAATCTTGTGCAAGGGTGGAGATTGATGG 1583

Qy 1657 GAGACTCCGGTGACAAACCCCTTAAGCGCAATAAATAGCTATATCTTCTTATACCATGGCAA 1716
Db 1584 GAGACTCCGGTGACAAACCCCTTAAGCGCAATAAATAGCTATATCTTCTTATACCATGGCAA 1643
Qy 1717 TATGTGGCATGCTCTCTGGATTCAATTCGTCGCAAAAGAGGTGAACAGAAAGGCGAAGAAA 1776
Db 1644 TATGTGGCATGCTCTCTGGATTCAATTCGTCGCAAAAGAGGTGAACAGAAAGGCGAAGAAA 1703
Qy 1777 TGGAGAGCTGACATGAGCTTAATGACAGACTCCAAGAGGGAATTCGAATGGACAGTTACA 1836
Db 1704 TGGAGAGCTGACATGAGCTTAATGACAGACTCCAAGAGGGAATTCGAATGGACAGTTACA 1763
Qy 1837 CCAGTTACTGCAATGCTGTGCTGACCTTCACTCAGCATCTCAGATAGACATGAGTGTCA 1896
Db 1764 CCAGTTACTGCAATGCTGTGCTGACCTTCACTCAGCATCTCAGATAGACATGAGTGTCA 1823
Qy 1897 AGGCAGAGATGGGTCTAGGTGACAGAAAGAGGAATTAATGGCTCTCTAGAGAAATGGTATG 1956
Db 1824 AGGCAGAGATGGGTCTAGGTGACAGAAAGAGGAATTAATGGCTCTCTAGAGAAATGGTATG 1883
Qy 1957 ACCAGATAGGCTGAAGTCT 2016
Db 1884 ACCAGATAGGCTGAAGTCT 1943
Qy 2017 TTGGGTCTATTCCGCCCATGGTGCAATGACGTAAGCAATGCCAATGGGCCCTCTGGTTGCTT 2076
Db 1944 TTGGGTCTATTCCGCCCATGGTGCAATGACGTAAGCAATGCCAATGGGCCCTCTGGTTGCTT 2003
Qy 2077 TATATTTGGTTTATGACACAGGAGATGTTTCTTCAAAAGTGCAACACCAATATGGCTTC 2136
Db 2004 TATATTTGGTTTATGACACAGGAGATGTTTCTTCAAAAGTGCAACACCAATATGGCTTC 2063
Qy 2137 TACTCTATGGTGGT 2196
Db 2064 TACTCTATGGTGGT 2123
Qy 2197 AGACATGGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTTCAGTATTTGAACCTGG 2256
Db 2124 AGACATGGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTTCAGTATTTGAACCTGG 2183
Qy 2257 CATCTGCCCTCACTGTGTGATTGATCAAAATATTGGGCTTCCCATCAGTACAACACATTT 2316
Db 2184 CATCTGCCCTCACTGTGTGATTGATCAAAATATTGGGCTTCCCATCAGTACAACACATTT 2243
Qy 2317 GTAAAGTGGGCTCTGT 2376
Db 2244 GTAAAGTGGGCTCTGT 2303
Qy 2377 GTCTCTTTCTGTAACATTTTATGGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2436
Db 2304 GTCTCTTTCTGTAACATTTTATGGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2363
Qy 2437 GTGCTGCCATCATGGCAATCTTTCAGATATGTCATCTCAGAAATGTGAAGCTGTTGAGAT 2496
Db 2364 GTGCTGCCATCATGGCAATCTTTCAGATATGTCATCTCAGAAATGTGAAGCTGTTGAGAT 2423
Qy 2497 TAAATTTGTGTCAATGTTTGGGACCATCTTAGGTATTTCTGCTCCCTCCCTGAAGATGATT 2556
Db 2424 TAAATTTGTGTCAATGTTTGGGACCATCTTAGGTATTTCTGCTCCCTCCCTGAAGATGATT 2483
Qy 2557 ACAGTGTTAACAGAGACTGCAAGAGTCTTTTATTTTGGGAGCCAGAGAGGGAAGTGT 2616
Db 2484 ACAGTGTTAACAGAGACTGCAAGAGTCTTTTATTTTGGGAG - CAGAGAGGGAAGTGT 2542
Qy 2617 TACTTGTGTATAAATGCTTTTGTGCTAAATATGAATGTCTCAAAATTAGCTGTGTGTA 2676
Db 2543 TACTTGTGTATAAATGCTTTTGTGCTAAATATGAATGTCTCAAAATTAGCTGTGTGTA 2602
Qy 2677 ATAGCCCGGGTTCACCTGCTGCTGAGGTCCCTTTTCTTCTGCGGCTGTGAATTCCT 2736
Db 2603 ATAGCCCGGGTTCACCTGCTGCTGAGGTCCCTTTTCTTCTTCTGCGGCTGTGAATTCCT 2662

QY 2737 GTACATATTTCTCTACTTTTTTGTATCAGGCTTCAATTCATATATGTTTAAATGTTGTC 2796
Db 2663 GTACATATTTCTCTACTTTTTTGTATCAGGCTTCAATTCATATATGTTTAAATGTTGTC 2722
QY 2797 TGAAGATGACTGTGATTTTTTTTCTTTTTTAAACCATGAAGACCGCTTTGACAGAG 2856
Db 2723 TGAAGATGACTGTGATTTTTTTTCTTTTTTAAACCATGAAGACCGCTTTGACAGAG 2782
QY 2857 CATGCTCTGCTGTTGTTTCAACAGCTTCTGCCCTCACATGACAGGATTTAAACAAC 2916
Db 2783 CATGCTCTGCTGTTGTTTCAACAGCTTCTGCCCTCACATGACAGGATTTAAACAAC 2842
QY 2917 AAAAATAATACTACAACCTCCCTGTAGTCTCTATATAAGTAGAGTCCCTTGACTCTG 2976
Db 2843 AAAAATAATACTACAACCTCCCTGTAGTCTCTATATAAGTAGAGTCCCTTGACTCTG 2902
QY 2977 CCTCTCTGTCAGTAGTGGCAGGATCTATTCGCATATTTCCGGAGCTCTTAGAGGGATGAG 3036
Db 2903 CCTCTCTGTCAGTAGTGGCAGGATCTATTCGCATATTTCCGGAGCTCTTAGAGGGATGAG 2962
QY 3037 GTTCTTTGAAACACAGTGAATAATTTAAATTAGTAACTTTTTCGAAGCAGTTTATTGACTG 3096
Db 2963 GTTCTTTGAAACACAGTGAATAATTTAAATTAGTAACTTTTTCGAAGCAGTTTATTGACTG 3022
QY 3097 TTATTGCTAAGAAGAAGTAAGAAAGAAAGCCCTGTTGGCAATCTTGTTATTCTTTAA 3156
Db 3023 TTATTGCTAAGAAGAAGTAAGAAAGAAAGCCCTGTTGGCAATCTTGTTATTCTTTAA 3082
QY 3157 GATTCTCGCAGTGGGATGGAATGAATGAATGGAATGGAATGGAATGGAATGGAATGGAAT 3216
Db 3083 GATTCTCGCAGTGGGATGGAATGAATGAATGGAATGGAATGGAATGGAATGGAATGGAAT 3142
QY 3217 GGGACAGCTTCCATGTTCTATTTGCTTACCTCTTAACCTGAATAAAAAAGCCTACAGTTT 3276
Db 3143 GGGACAGCTTCCATGTTCTATTTGCTTACCTCTTAACCTGAATAAAAAAGCCTACAGTTT 3202
QY 3277 TAGAAAAA 3285
Db 3203 TAGAAAAA 3211

RESULT 3

US-07-674-287B-3

; Sequence 3, Application US/07674287B

; Patent No. 5414076

; GENERAL INFORMATION:

; APPLICANT: Bryan Mark O'Hara

; TITLE OF INVENTION: Gibbon Ape Leukemia

; TITLE OF INVENTION: Virus Receptor

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. Karen A. Lowney

; ADDRESSEE: American Cyanamid Company

; STREET: 1937 West Main Street

; STREET: P.O. Box 60

; CITY: Stamford

; STATE: CT

; COUNTRY: USA

; ZIP: 06904-0060

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC AT

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: ASCII converted from IBM DW4

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/674,287B

; FILING DATE: 19910325

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Lowney, Karen A., Dr.

REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 31,104-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203 321 2361
TELEFAX: 203 321 2971
TELEX: 710 474 4059
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3260 Bps 681 Amino Acid Residues
TYPE: NUCLEOTIDE AND AMINO ACID SEQUENCES
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: DNA and Protein
US-07-674-287B-3

Query Match 64.8%; Score 2133; DB 1; Length 3260;

Best Local Similarity 84.5%; Pred. No. 0;

Matches 2662; Conservative 0; Mismatches 420; Indels 67; Gaps 21;

QY 162 CCTCCCTTTTCCCTGGATGAACCTTGCCTCTTTCTCTCTCGCCATGGAATTCGTCTC 221
Db 157 CACTTCGTCCTCCCGCAGGATGAACCTTGCCTCTTTCTCTAAATCGCCATGGAATTCGTCTC 216
QY 222 CGTCTTTTAGCCCTCTCGAGCAAGAAACCCAGACAAACAGATGCCATAC---GCAG 278
Db 217 CGTCTTTTAGCCCTCTCGAGCAAGAAACCCAGACAAACAGATGCCATAC---GCAG 276
QY 279 CGTATAGCAGTAACCTCCCGAGCTCGGTTTCTGTCGCTAGTTTACAGTATTTAAATTTAT 338
Db 277 CGTATAGCAGTAACCTCCCGAGCTCGGTTTCTGTCGCTAGTTTACAGTATTTAAATTTAT 336
QY 339 ATAATATATATTTATTTATATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTTGA 398
Db 337 ATAATATATATTTATTTATATAGCATTTT---GATACCTCATTTCCGTTTACACATCTCAA 393
QY 399 AAGCGCTCAGTAGTCTCTTA-----CTAAACCAACCTACTTCCAGAGAA----- 444
Db 394 AAGCGCTTAGTAATTTCTTATTTATTAAGAACCACTACACTAGAGAATGGAATCTAC 453
QY 445 --TGGCAACGGTGATTACCAGTACTACAGCTCTACCGCGCTTCTGCTCTTGTGGA 502
Db 454 TGTGGCAACGATTACTAGTACCTAGCTGCTGTTACTGCTTCCGCTCCACCGAAGTATGA 513
QY 503 CTACCTATGATGCTCATCTCGGCTTCAATTTGCAATTTGCTTGGCATTTCTCGTGGG 562
Db 514 CAATCTATGATGCTCATCTCGGCTTCAATTTGCAATTTGCTTGGCATTTCTCGTGGG 573
QY 563 AGCAATGATGTAGCAAAATTTCTTTGGTACAGCTGTGGCTCAGGTGTAGTACCCTGAA 622
Db 574 AGCAATGATGTAGCAAAATTTCTTTGGTACAGCTGTGGCTCAGGTGTAGTACCCTGAA 633
QY 623 GCAAGCTCCTCATCTAGCTAGCATCTTTGAAACAGTGGCTCTGCTTACTGGGGCCAA 682
Db 634 GCAAGCTCCTCATCTAGCTAGCATCTTTGAAACAGTGGCTCTGCTTACTGGGGCCAA 693
QY 683 AGTAGCGAAACCATCCGGAAGGCTTGTATGACGTGGAGATGTACAACTCGACTCAAGG 742
Db 694 AGTAGCGAAACCATCCGGAACGGCTTGTATAGATGTGGAGCTGTACAAACGAACTCAAG 753
QY 743 GCTACTGATGGCGGCTCAGTCAGTGTCTATGTTTGGTCTGCTGTGTGTGCAACTCTGTGC 802
Db 754 TCTGCTCATGGCTGGCTCGCTCAGTGTCTATGTTTGGTCTGCTGTGTGTGCGAGCTCTGTC 813
QY 803 TTCTGTTTTGAAGCTCCCTTATTTCTGGAACCCATTTGTTTGGTGGCAACTATTGTTT 862
Db 814 TTCTGTTTTGAAGCTTCCGATTTCTGGGACCCATTTGTTTGGTGGCAACCACTTGGTTT 873
QY 863 CTCCCTCGTGGCAAAAGGGGAGGAGGTGTCAAGTGGTCTGAATCTGATAAATAATTTGTAT 922
Db 874 CTCCCTCGTGGCAAAAGGGGAGGAGGTGTCAAGTGGTCTGAATCTGATAAATAATTTGTAT 933
QY 923 GTCTTGGTTCTGTCCTCCCACTCTCTTCTGGAATTTATGCTGGAATTTTATTTCTCTGTT 982

Db 934 GTGCTGGTTCCTCTCCGCTGCTTTCTGGTATATATGCTGGAATTTATTTCTTCTCTGT 993
Qy 983 TCGTGCAATTCATCTCCATAGGCAGATCCAGTTCCTAATGGTTTGGAGCTTTGGCAGT 1042
Db 994 TCGTGGCTTCATCTCCGTAAAGCAGATCCGGTTCCTAATGGTTCAGAGCTTTTACCAAT 1053
Qy 1043 TTTCTATGCTGCACAGCTGGAAATAAACCTCTTTTCCATCATGTATATCTGGAGCAGCGTT 1102
Db 1054 TTTTATGCTGCACAAATCGAATCAACCTCTTTTCCATTTATGATATCTGGAGCAGCGTT 1113
Qy 1103 GCTGGGCTTTGACAACTTCCTCTGTTGGGTACCATCTCATCTCGGTGGATGTGCAGT 1162
Db 1114 GCTGGGCTTTGACAACTTCCTCTGTTGGGTACCATCTCATCTCGGTGGATGTGCAGT 1173
Qy 1163 TTTCTGTGCTTATCTGCTGGTCTTTGTATGTCTCCAGATCAAGAGAAAATTTGAACG 1222
Db 1174 TTTCTGTGCTTATCTGCTGGTCTTTGTATGTCTCCAGATCAAGAGAAAATTTGAACG 1233
Qy 1223 AGAAATAAGTGTAGTCTCTGAAAGCCCTTTAATGGAATAAGAAATAGCTTGAAGA 1282
Db 1234 AGAATTAAGTGTAGTCTGAAAGCCCTTTAATGGAATAAGAAATAGCTTGAAGA 1293
Qy 1283 AGACCATGAAGAAACAAAGTTGTCTGTTGTGTATTTGAAACAAAGCATCTCTGTCTGA 1342
Db 1294 AGACCATGAAGAAACAAAGTGTCTGTTGTGTATTTGAGCATAGGAATCTCTGTCTGA 1353
Qy 1343 GGTAGGCTTCCACTGTGCTCCCTCCAGGCTGTGGTGGAGAGACAGTCTCATTCAA 1402
Db 1354 GGTAGTGTGCTTCCACTGTGCTCCCTCCAGGCTGTGGTGGAGAGACAGTCTCATTCAA 1413
Qy 1403 ACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGGCTTCCAGCGTGGACTTGAAGA 1462
Db 1414 ACTTGGTGTGCTGGAGAGGCTCCGAGCGAGAGCGCTTCCCA---TGGACCTGAAGA 1470
Qy 1463 GGAACACAGCATAGATAGCACCGTGAATGGTGCAGTGCAGTTGCCTAATGGGAACCTTGT 1522
Db 1471 GGAGCCAGCATAGACAGCACCATCAATGTGCAGTGCAGTTGCCTAATGGGAACCTTGT 1530
Qy 1523 CCAGTTCAAGCTGACCAACCAATAACTCCAGTGGCCCTACCACTATCACAC 1582
Db 1531 TCAGTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1590
Qy 1583 CGTGCAATAGGATTCGGGCTGTACAAAGAGCTACTCCATAAATTTACATCTTGCCAAAGT 1642
Db 1591 CGTGCAAGATTTCTGGCTGTACAGAGAGCTCTCATTAAGTTACATCTGSCCAAGT 1650
Qy 1643 GGGAGATTGATGGGAGACTCCGGTGACAAACCCCTTAAGGCGCAATTAAGTATATCTTC 1702
Db 1651 GGGAGACTGATGGGAGATTCGGGGAACAGCCCTTGAGACGCAACAAACAGCTACACTTC 1710
Qy 1703 CTATACCATGGCATATGGCATGCTCTGGATTCATTCGGTCCCAAGAGGTGAACA 1762
Db 1711 CTACATATGGCAATATGTGGCATGCTCCCTGGATTCATTCGGTCCCAAGAGGTGAACA 1770
Qy 1763 GAAGGCGAAGAAATGGAGAGCTGACATGGCCCTTAATGCAGACTCCAAAGAGCGAATTCG 1822
Db 1771 AAAGGAGATGAATGGAGAGCTGACATGGCCCTTAATGCAGATACCAAGAGCGGATTCG 1830
Qy 1823 AATGGACAGTTACACAGTTACTGCAATGTGTGTGCTGACCTTCACCTCAGCATCTGAGAT 1882
Db 1831 AATGGACAGTTACACAGTTACTGCAATGTGTGTGCTGACCTTCACCTCAGCATCTGAGAT 1890
Qy 1883 AGACATGAGTCAAGCAGAGATGGGTCTAGGTGACAGAAAAGGATTAATGGCTCTCT 1942
Db 1891 GGACATGAGTGAAGCTGAGATGGGCTTGGGTGACAGAAAAGGATTAATGGCTCTCT 1950
Qy 1943 AGAAGATGGTATGACCAAGATAGCTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2002
Db 1951 TGAAGATGGTATGACCAAGATAGCTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2010
Qy 2003 CTTTACAGCTGCTTTGGGCTATTCGCCCATGGTGGCAATGAGTAAAGCAATGCCATTTGG 2062
Db 2011 CTTTACAGCTGCTTTGGGCTATTTGCCCATGGTGGCAATGAGTAAAGCAATGCCATTTGG 2070

Qy 2063 GCTCTGGTTCCTTATATTTGTTTATGACACAGAGATGTTTCTTCAAAAGTGCAAC 2122
Db 2071 CCTCTGGTTCCTTATGTTATCTGTTTAT---AAACAAGAGCTCTACAAAAGCGCAAC 2127
Qy 2123 ACCAATATGGCTTCTACTCTATGGTGGTGTGGTATCTGTGTGTGTGTGTGTGTGTGGG 2182
Db 2128 ACCCATATGGCTTCTGCTTTATGGTGGTGTGGCATTTGCATGGGCTGTGTGTGTGGG 2187
Qy 2183 AAGAGAGTTATTCAGACCATGGGAAGGATCTGACACCGATCAGACCTCTAGTGGCTT 2242
Db 2188 AAGAAGAGTTATTCAGACCATGGGAAGGATCTGACACCGATCAGACCTCTAGTGGCTT 2247
Qy 2243 CAGTATTAAGTGGCATCTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2302
Db 2248 CAGTATTAAGTGGCATCTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2307
Qy 2303 CAGTACAAACATTTGTAAGTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2362
Db 2308 CAGCACAACATTTGTAAGTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2367
Qy 2363 GGT 2422
Db 2368 GGT 2427
Qy 2423 TTTCTGAGTTATCAGT 2482
Db 2428 CTTCTGGGTTATCAGT 2487
Qy 2483 AAGCTGT 2542
Db 2488 ACGCT---GGGTTGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2542
Qy 2543 CTTGAAGAAATGATACAGT 2602
Db 2543 CTTGAAGAAATGATACAGT 2596
Qy 2603 GAGGAGGAGT 2662
Db 2597 TGGGAAGGAGT 2656
Qy 2663 ATTAGCTGT 2722
Db 2657 ATTAGCTGT 2715
Qy 2723 GGT 2782
Db 2716 GGT 2774
Qy 2783 TTTAATGTTGTCTCTGAAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2842
Db 2775 TTTAATGTTGTCTGTGA-----GATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2826
Qy 2843 GCGGTTTGAAGAGATGCTGT 2902
Db 2827 GAGCCATTTGATGGCATGT 2886
Qy 2903 AGGGATTTAACAACAAATATACTAACAATTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2962
Db 2887 AGGGATTTAACAACAAATATACTAACAATTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2946
Qy 2963 TCTTGT 3022
Db 2947 AC---GGCACTCTGTCTC---CTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
Qy 3023 CTTAGAGGATGAGTGTCTTTGAACACAGTGAATAATTTAAAT---TAGTAACCTTTTTCG 3079
Db 3001 CTTAGAGGATGAGTGTCTTTGAACACAGTGAATAATTTAAAT---TAGTAACCTTTTTCG 3059
Qy 3080 AAGCAGTTTATGACGT 3139
Db 3060 AAGCAGTTTATGACGT 3112

QY 3140 CTTGGTATTCTTTAAGATTTCTGGCAGTGTGGGATGGAATGAATGGAAATGTGAA 3199
Db 3113 TGTGGTATTCTTCTGAGATTTCTGGCAGTGTGGGATGGAATGAATGGAAATGTGAA 3172
QY 3200 CTTTGGGCAAGTTAAATGGGACAGCCTTCCATGTTCAATTTGTCTACCTCTTAAGTAATA 3259
Db 3173 CTTTGGGCAATTAATGGGACAGCCTTCCATGTTCAATGTTCTGTCTACCTCTTAAGTAATA 3231
QY 3260 AAAAGCCTACAGTTTTTAGAAAAAAA 3288
Db 3232 AAAAGCCTACAGTTTTTAGAAAAAAA 3260

RESULT 4

US-08-436-900A-3
; Sequence 3, Application US/08436900A
; Patent No. 5874264
; GENERAL INFORMATION:
; APPLICANT: O'Hara, Bryan M.
; TITLE OF INVENTION: Gibbon Ape Leukemia Virus Receptors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,900A
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,104-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-683-2158
; TELEFAX: 201-683-4117
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3260 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 443..2488
; US-08-436-900A-3

Query Match 64.8%; Score 2131.4; DB 2; Length 3260;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2661; Conservative 0; Mismatches 421; Indels 67; Gaps 21;

QY 162 CCTCCCTTTTCCCTGATGAATCTGGTCTTCTCTTCTCCGCCATGGAATTCGCTC 221
Db 157 CACTTCGTCCTCCCGAGATGAATCTGGTCTTCTCTAATCCGCCATGGAATTCGCTC 216
QY 222 CGTGTCTTTAGCCCTCTGAGCCCAAGAAACCCAGACAACAGATGCCCATAC---GCAG 278
Db 217 CGTGTCTTTAGCCCTCTGAGCCCAAGAAACCCAGACAACAGATGCCCATGAGCAG 276
QY 279 CGTATAGCAGTAATCCCGCAGCTCGGTTTCTGTGCGGTAGTTTACAGTATTTAATTTAT 338
Db 277 CGTATAGCAGTAATCCCGCAGCTCGGTTTCTGTGCGGTAGTTTACAGTATTTAATTTAT 336
QY 339 ATAATATATATTTTATTATAGCATTTTGTGATACCTCATATTTCTGTTTACACATCTTGA 398

Db 337 ATAATATATATTTTATTATATAGCATTTT---GATACCTCATTTCCGTTTACACATCTCAA 393
QY 399 AAGGCGCTCAGTAGTCTCTCTTA---CTAAACAACCACTACTCCACAGAA-----444
Db 394 AAGCCCTTAGTAATTTCTCTTATTTAAAGAACCACTACACTAGAGAATGAAATCTAC 453
QY 445 --TGGCAACGGTGATTACACAGTACTACAGCTGTACCGCGCTTCTGGTCTTGTGGTGA 502
Db 454 TGTGGCAACGATTACTAGTACCTAGCTGCTGTACTGCTCCGCTCCACCGAAGTATGA 513
QY 503 CTACCTATGATGCTCATCTCGGCTTCATTTATTTGATTTGTTTGGCATTTCTCCGTGGG 562
Db 514 CAATCTATGGATGCTCATCTCGGCTTCATTTGATTTGCTTGGCATTTGTTTGGCATTTCTCCGTGGG 573
QY 563 AGCCATATGATGTAGCAAAATTTCTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCTGAA 622
Db 574 AGCCATATGATGTAGCAAAATTTGTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCTGAA 633
QY 623 GCAAGCCTGCATCTAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCCAA 682
Db 634 GCAAGCCTGCATCTAGCTAGCATCTTCGAAACTGTGGGCTCCGCTTGTGGGGGCCAA 693
QY 683 AGTGAGCGAAACCATCCGGAAGGGCTTGATTGACGTGGAGATGTACAACTGACTCAAGG 742
Db 694 AGTGAGCGAAACCATCCGGAACGGCTTGATAGATGTGGAGCTGTACAACTCAAGA 753
QY 743 GCTACTGATGGCGGCTCAGTGTGCTATGTTGGTCTGTGCTGTGGCAACTCTCGTGGC 802
Db 754 TCTGCTCATGGCTGGCTCCGTCAGTGTATGTTGGTCTGTGCTGTGGCAGCTCGTGGC 813
QY 803 TTCGTTTTTGAAGCTCCCTATTTTGGAAACCACTGTATTTGTTGGTGGCAACTATTGGTTT 862
Db 814 TTCGTTTTTGAAGCTTCCGATTTCTGGGACCACTGTATTTGTCGGTGGCAACCACTGGTTT 873
QY 863 CTCCTCTGTGGCAAAAGGGCGAGAGGTGTCAAGTGGTCTGAATCTGATATAAATTTGTAT 922
Db 874 CTCCTCTGTGGCAAAAGGGCGAGAGGTGTCAAGTGGTCTGAATCTGATATAAATTTGTAT 933
QY 923 GTCTGTTGTTGTTGCTCCCACTGCTTCTGGAATTTATGCTGGAATTTTATTTCTCTCGT 982
Db 934 GTCTGTTGTTGTTGCTCTCCGCTCTCTGGAATTTATGCTGGAATTTTATTTCTCTCTGT 993
QY 983 TCGTGATTCATCTCCATTAAGGCAGATCCAGTTCCTAATGTTTCCGAGCTTTGCGCAGT 1042
Db 994 TCGTGGTTCATCTCCGTTAAGGCAGATCCGTTCTAATGCTTACGAGCTTTTACCAAT 1053
QY 1043 TTTCTATGCTGCACAGTTGGAAATAAACTCTTTTCCATCATGTATATCTGGAGCACGTT 1102
Db 1054 TTTTATGCTGCACAAATCGGAATCAACCTCTTTTCCATTTATGTATATCTGGAGCACGTT 1113
QY 1103 GCTGGGCTTTGACAAACTCTCTGTGGGTACCATCTCATCTCGTGGGTGTGTCAGT 1162
Db 1114 GCTGGGCTTTGACAAACTCTCTGTGGGTACCATCTCATCTCGTGGGTGTGTCAGT 1173
QY 1163 TTTCTGTGCCCTTATCGTCTGTTCTTTGTATGTCCAGATGAAGAGAAAAATTTGAACG 1222
Db 1174 TTTCTGTGCCCTTATCGTCTGTTCTTTGTATGTCCAGATGAAGAGAAAAATTTGAACG 1233
QY 1223 AGAAATAAAGTGTAGTCTCTTGAAGCCCTTAATGGAAAAAAGAAATAGCTTTGAAAGA 1282
Db 1234 AGAAGTAAAGTGTAGTCTCTGAAAGTCCCTTAATGGAAAAAAGAGAGCAACTTAAAGA 1293
QY 1283 AGACCATGAAGAAACAAAGTTGTTGTTGTATATTGAAACCAAGCATCTCTGTTTCTGA 1342
Db 1294 AGACCATGAAGAAACAAAGATGGCTCTCGGAGACGTTGAGCATAGGAATCTCTGTCTGA 1353
QY 1343 GGTAGGGCTGCGACCTGTGCCCTCCAGGCTGTGGTGGAGAGAGAAAGCTCTCATTTCAA 1402
Db 1354 GGTAGTGTGCGACCTGGGCCACTCGGGCTGTGGTGGAGAGAGAGCGGTGTCTATTCAA 1413
QY 1403 ACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGAGGCTTCCAGCGGTGTGAAAGA 1462

1414 ACTTGTGACCTGGAGGAGGCTCCGGAGCGAGAGCGGCTTCCCA---TGGACCTGAAGGA 1470
1463 GGAACACAGATAGATAGACACCTGATGTCAGTCAGTCAGTTCCTTAATGGGAACCTTGT 1522
1471 GGAGACAGATAGACAGACCACTCAATGTCAGTCAGTTCCTTAATGGGAACCTTGT 1530
1523 CCAAGTTCAGTCAAGCCGTGACGAACCAATAAATAAATCCAGTGGCCACTTACCAGTATCACAC 1582
1531 TCAGTTCAGTCAAACTGTGAGCAACGATCAACTCCAGTGGCCACTTACAGTATCACAC 1590
1583 CGTGCAATAGGATTCGGCCCTGTACAAAGAGCTACTCCATAAATACATCTTCCCAAGGT 1642
1591 CGTGCAAGGATTCCTGGCTGTGACAAAGAGCTGCTCCATAAGTTACATCTGCGCAAGGT 1650
1643 GGGAGATTGATGGGAGACTCCGGTGCACAAACCTTTAAGCGGCAATAATAGCTATACTTC 1702
1651 GGGAGACTGATGGGAGATTCTGGGACAAAGCCCTTGAGACGCAACCAAGCTTACACTTC 1710
1703 CTATACCATGGCAATATGTGGCATGCTCTGGATTCAATCCGTGCCAAAGAGGTGAACA 1762
1711 CTACATATGGCAATATGTGGCATGCTCTGGATTCAATCCGTGCCAAAGAGGTGAACA 1770
1763 GAAGGGCAAGAAATGGAGAGCTGACATGGCCCTAATGCAAGTCCAAAGAGCGAATTCG 1822
1771 AAAGGGAGATGAATGGAGAGCTGACATGGCCCTAATGCAAGTCCAAAGAGCGAATTCG 1830
1823 AATGGACAGTTACACAGTTACTGCAATGCTGTGTCGACTTCACTCAGCATCTGAGAT 1882
1831 AATGGACAGTTACACAGTTACTGCAATGCTGTGTCGACTTCACTCAGCATCTGAGAT 1890
1883 AGACATGAGTGTCAAGGACAGATGGGTCTAGGTGACAGAAAGAGGAGTAAATGGCTCTCT 1942
1891 GGACATGAGTGTGAGGCTGAGATGGGCTGGGTGACAGAAAGAGGAGTAAATGGCTCTCT 1950
1943 AGAAGATGATGACACAGATGAAGCTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2002
1951 TGAAGATGATGACACAGATGAAGCTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2010
2003 CTTACAGCTGCTTGGGTCAATCGCCATGGTGGCAATGAGTAAGCAATGCCATTGG 2062
2011 CTTACAGCTGCTTGGGTCAATCGCCATGGTGGCAATGAGTACAGCAATGCCATTGG 2070
2063 GCCTCTGCTTGTATATATGTTTATGACACAGAGGATGTTTCTTCAAAAGTGGCAAC 2122
2071 CCCTCTGCTTGTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 2127
2123 ACCAATATGCTTCTACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2182
2128 ACCCATATGCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2187
2183 AAGAGAGTTATCCAGACCATGGGAGGATCTGACACCGATCACACCTCTAGTGGCTT 2242
2188 AAGAGAGTTATCCAGACCATGGGAGGATCTGACACCGATCACACCTCTAGTGGCTT 2247
2243 CAGTATTGAATGGCATCTGCCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2302
2248 CAGTATTGAATGGCATCTGCCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2307
2303 CAGTACAAACATGTTAAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2362
2308 CAGCACAACATGTTCAAAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2367
2363 GGCTGTTGATGGGCTCTCTTCTGATACATTTTATGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2422
2368 GGCTGTTGATGGGCTCTCTTCTGATACATTTTATGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2427
2423 TTCTGGAGTTATCAGTCTGCCATCATGGCAATCTTCAGATATGTCATCTCAGAAATGTG 2482
2428 CTCTGGGGTTATCAGTCTGCCATCATGGCAATCTTCAGATATGTCATCTCAGAAATGTG 2487
2483 AAGCTGTTGAGATTAATTTGTCATATGTTGGGACCATCTTGGGATTTCTGCTGCTC 2542
2488 ACGCT----GGGGTTGAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2542

2543 CCTGAAGATGATTACAGTGTAAACAGAGACTGACAGAGTCTTTTATTTTGGAGCCA 2602
2543 CTAGAGAACGCTCAGTG-TGGCTGAAGACAGGCAAGGGTCT-----TAAGAGCGCG 2596
2603 GAGGAGGGAAGTGTACTTTGCTATAACTGCTTTTGTCTAAATATGAATGTCTCAAA 2662
2597 TGGGAAGGAAGTGTAAATTTACACTATAATTTGCTTTTGTCTAAATATGAATGTCTCAAA 2656
2663 ATTAGCTGTGTAATAATAGCCCGGTTCCACTGGCTCTGCTGAGGTCCTTCTTCTG 2722
2657 ATTAGCTGTGTAATAATAGCCCGGTTCCACTGGCTTTCAATTTCAATTTCAAGGT-CCCTTTTCTCTG 2715
2723 GCTGTGAATTCCTGTACATAATTTCTACTTTTGTATCAGCTTCAATTTCAATTTATCT 2782
2716 GCTGTGAATTCCTGTACATAATTTCTCTAC-TTTTGTATCAGGCTCAATTTCAATTTATCT 2774
2783 TTTAATGTTGCTCTCTGAAGATGACTTGTGATTTTTTTTTTTTTTTTTTAAACCATGAAGA 2842
2775 TTTAATGTTGCTCTGTA-----GATAACTTAGTGGTCTTTTTTAAACAGCCAGCA 2826
2843 GCGCTTTGACAGAGCATGCTCTGCTGTTGTTGTTTCAACAGCTTCTGCCCTCACATGCAC 2902
2827 GAGCCATTTGATGGCATGTACTGCTTGTGCGCTCACAGCTTCTTCCCAACATGCAC 2886
2903 AGGGATTTAACAACAATAATACTACAACCTTCCCTTGTAGTCTCTTATATAAGTAGAG 2962
2887 AGGGATTTAACAACAATAATACTACAACCTTCCCTTGTAGTCTCTTATATAAGTAGAG 2946
2963 TCTTTGTTGCTCTGCCCTCTCTGCTAGTGTGCGAGTCTATTGGCATATTCGGGAGCTT 3022
2947 AC---GGCACTCTGCTC-CCTGCTACTAGTGGCAGTCTCTGTTG--ATGTTGACAACTT 3000
3023 CTTAGAGGATGAGGTTCTTTGAAACACAGTGAATTTAAAT---TAGTAACTTTTTCG 3079
3001 CTTAGAGGCGGAGAACTTTTG-GCACAGTGGAAATATAAGTTTGTAGTAACTCTTTCG 3059
3080 AAGCAGTTTATTTGACTGTTTATTTGCTAAGAAAGTAAAGAAAGAAAGCTGTGGCAAT 3139
3060 AAACAGTTTCCAGC-ATGTTGCTAAGAAAGCGGAGACAAA-----GCCCTGGCGT 3112
3140 CTTGTTTATTTCTTTAAGATTTCTGCGAGTGTGGATGGATGAATGAAGTGAATGTGAA 3199
3113 TGTGTTTATTTCTCTGAGATTTCTGCGAGTGTGGATGGTGAATGAAGTGAATGTGAA 3172
3200 CTTTGGGCAAGTAAATGGACAGCTTCCATGTTCAATTTGCTTACCTCTTAACCTGAATA 3259
3173 CTTTGGGCAATTTCAATGGACAGCTTCCATGTTTCACTGCTACCTCTTAACCTGAAT- 3231
3260 AAAAAAGCTTACAGTTTATTAAGAAAAAAA 3288
3232 AAAAAAGCTTACAGTTTATTAAGAAAAAAA 3260

RESULT 5

US-08-050-684-1
; Sequence 1, Application US/08050684
; Patent No. 5550221
; GENERAL INFORMATION:
; APPLICANT: Johann Dr., Stephen V.
; APPLICANT: Van Zeijl Dr., Marja
; APPLICANT: O'Hara Dr., Bryan M.
; TITLE OF INVENTION: Amphotropic Virus Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: 1937 West Main Street
; CITY: Stamford
; STATE: CT
; COUNTRY: United States of America
; ZIP: 06904-0060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/050,684
FILING DATE: 16-APR-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 31937-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-321-2361
TELEFAX: 203-321-2971
TELEX: 710-474-4059
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3175 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 244..2202
US-08-050-684-1

Query Match 18.8%; Score 617.2; DB 1; Length 3175;
Best Local Similarity 60.3%; Pred. No. 2.6e-163;
Matches 1203; Conservative 0; Mismatches 718; Indels 75; Gaps 8;
QY 496 TGGTGACTACCTATGATGCTCATCTGGGCTTCATTATGCTATGCTTGGCATTTCTTGGCATTTCT 555
DB 251 TGGATGAGTATTGTGGATGGTCAATTTGGTTTCAATAGTCTTCACTTGGCCCTTTT 310
QY 556 CCGTGGGAGCAATGATGAGCAAAATCTTTTGGTACAGCTGFGGCTCAGGTGTAGTGA 615
DB 311 CTGTGGTGCACACGATGTGCCAATCTCTTGGTACAGCCGTGGCTCTGGTGTGGTGA 370
QY 616 CCTGAAGCAAGCTGCATCTAGCTAGCATCTTTGAAACAGTGGGCTCTGCTTACTGG 675
DB 371 CCTTGAAGGAGGAGCATGCAATTTAGCTTCAATATTTGAAACCAACCGGCTCCGTGTACTAG 430
QY 676 GGGCCAAAGTGAGCGAAACCATCCGGAAGGCTTGATGACGTGGAGATGTACAACTCGA 735
DB 431 GCGCCAAAGTAGAGAAACCATTCGCAAGGATATCAATGACGTGAACCTGTACACAGAGA 490
QY 736 CTCGAAGGCTACTGATGGCGGCTCAGTCAGTGCTATGTTTGGTTCGTGTGTGGCAAC 795
DB 491 CGGTGGAGACTCTCATGGCTGGGGAAGTTAGTGCCATGTTGGTTCCGCTGTGTGGCAGC 550
QY 796 TCGTGCTTGGTTTGAAGCTCCCTATTTCTGGAAACCATTTGATTTGTTGGTGAACATA 855
DB 551 TGATGCTTCTTCTTGGAGCTTCCAATCTCAGGAACGCACATGCAATTTGGGTTCTACTA 610
QY 856 TTGGTTTCTCCCTCGTGGCAAGGGGCGAGAGGCTCAAGTGTCTGAACTGATATAAAA 915
DB 611 TAGGATTTCTACGTGTGCAATCGGTACCAAGGTTGTGAGTGTGGAGCTTGTCAAGA 670
QY 916 TTGTGATGCTTGGTTTGGTGTCCCACTGCTTTCTCGAAATTAATGCTGGAATTTTATCT 975
DB 671 TTGTGCTTCTTGGTTTATATCTCCACTGTGTGCTGTTTCAATGCTGCGCTGCTGTTG 730
QY 976 TCTGTGTTCTGTCATTCATCTCCATAGGCAGATCCAGTTCCTAATGTTTGGAGCTT 1035
DB 731 TACTCATGAATTTTCACTCTTAAAGAGGAGACCCCTGTTCCTCAATGGCCTCCGGGCAC 790
QY 1036 TGCAGTGTCTATGCTGCACAGTTTGGAAATAAACCTCTTTTCCATCATGTATCTACCTGGAG 1095
DB 791 TCCAGTATTCTATGCTGTACCATAGCAATCAATGTCTTTTCCATCATGTACACAGGAG 850

QY 1096 CACCGTTGCTGGGCTTTTGACAAATTCCTCTCTGTGGGTACCATCTCTCATCTCGTGGGAT 1155
DB 851 CACCACTGCTCGGCTTG---TTCTCCCATGTGGGCCATAGCCCTCATTTCTTTGGTG 907
QY 1156 GTGCAGTTTCTGTGCCCTTATCGTCTGTTCTTTGTATGTCCAGAGTGAAGAGAAAA 1215
DB 908 TCGCCCTCTGTTCGCTTTTGTGTGGCTCTTCTGTGTGTCGTGATGCGGAGAAAA 967
QY 1216 TTGAACGAGAAATAAGTGTAGTCTTCTGAAAGCCCTTAAATGGAAGAAAAAGATAGCT 1275
DB 968 TAACAGGCAATTAACAAAGAAAGGTGCTTTATCAGGAGTATCTGACGAAGGCTCAGTA 1027
QY 1276 TGAAGAAGACCATGAAGAAAAAAGTGTCTGTGTTGGTATATATGAAAAACAAGCATCTCT 1335
DB 1028 AGGTTTCAGGAAGCAGAGTCCCAAGTATT-----TAAAGAGCTACCAAG 1070
QY 1336 TTTCTGAGGTAGGCTGCCACTGTGCCCCCTCCAGGCTGTGGTGGAGGAGAGAACAGTCT 1395
DB 1071 TGCCAAAGGCTAATGATGACAGCACCATCCCGCTTCAGGGAGCAGCAGGGAG----- 1122
QY 1396 CATTCAAACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGAGAGGCTTCCAGCGTGGACT 1455
DB 1123 -----ACACTGGGACCTCGGAAGGCACTTCTGCG-----GGCAGCCACCTCGGGCT 1170
QY 1456 TGAAGAAGAAACCAAGCATAGATAGCACCGTGAATGGTGCAGTGCAGTGTGCCTTAATGGGA 1515
DB 1171 GCATACGGAAGAGCACTGCTCATGCCATGGCTCTGTGAATCGCCCATCTCCAACGGC 1230
QY 1516 ACCTTGTCCAGTTCACTCAAGCGCTCAGCAACCAATAAATCCAGTGGCCACTACAGT 1575
DB 1231 ACCTT-----CGGCTTCAGCGGCCACACAGGAGCAGCGTCTATGTGT 1273
QY 1576 ATCACACCTGTCATAGGATTCGGGCTGTACAAAGAGCTACTCCATAAATTAATCATCTTG 1635
DB 1274 ACCACACCTGTCAACAAAGACTCGGGCTCTACAAAGATCTGTGCAACAAATCCACATCG 1333
QY 1636 CCAAGTGGGAGATTCATGAGGAGACTCCGGTGACAA-----ACCTTAAAGCGCAATA 1689
DB 1334 ACAGGGGCCCGAGGAGAGCAGCCAGGAAAGCAATACCGGCTGCTCCGCGGAAACA 1393
QY 1690 ATAGCTATACCTTCTATACCATGGCAATATGTGGCATGCTCTGGATTCATTTCCTGGCCA 1749
DB 1394 ACAGTTACACTGCTACACCGCAGCCATTTGTGGCTGCCAGTGCAAGCCACCTTCGAG 1453
QY 1750 AAGAAGGTGAACAGAGGCGAAGAAATGGAAGCTGACATGGCCCTAATG---CAGACT 1806
DB 1454 CTGCGGACTCATCGGCCCCAGAGGACAGTGAAGCTGTGGGCGACACCGTCTCTACT 1513
QY 1807 CCAAGAAGCGAATTCGAATGGACAGTTACACCAAGTTACTGCAATGCTGTCTGACCTTC 1866
DB 1514 CCAAGAAGAGGCTGCGCTACGACAGCTACTCGAGCTACTGTAAACGCGGTGGCAGAGCGG 1573
QY 1867 ACTCAGCATCTAGA-----TAGACATGAGTGTCAAGCAGAGATGGGTCTAGGTG 1917
DB 1574 AGATCGAGCGGAGGAGGCGCGCTGGAGATGAAGCTGGCGTGGAGCTGGCGAGCCCTG 1633
QY 1918 ACAGAAAAAGGAAGTAAATGCTCTCTAGAAAGATGTTATCACCAAGGATAAGCCCTGAAGTCT 1977
DB 1634 ACAGCGCGCAGAGACCCCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1593
QY 1978 CTCTCTCTTCCAGTTCTTCAGAGATCTCTTACAGCTGCTTTTGGGTCAITTCGCCCATGGTG 2037
DB 1694 ACCTCTGTTCATTTCTGTCAGGCTCTCACCGCTGTTTGGGTCTCTTGTCTACCGCG 1753
QY 2038 GCAATGACGTAAAGCAATGCCATTTGGGCTCTGGTGTCTTATATTTGGTTTATGACACAG 2097
DB 1754 GCAATGACGTAAAGCAATGCCATTTGGGCTCTGGTGTCTTATTTGAGTGAATTTACAAACAAG 1813
QY 2098 GAGATGTTTCTTCAAAAGTGGCAACCAATATGCTTCTACTCTATGTTGGTGTGGTGA 2157
DB 1814 GCGGGGTAAAGCAAGAGAGAGTACACCGCTCTGGCTGCTGTTTATGAGAGGTTGGAA 1873
QY 2158 TCTGTGTTGCTGTGGGTTTGGGGAAGAGAGTTATCCAGACCATTTGGGGAAGGATCTGA 2217

Db 1874 TCTGACAGGCTCTGGTCTGGGGGAGAAGAGTGATCCAGACCATGGGGAAGGACCTCA 1933
Qy 2218 CACCGATACACCTCTAGTGGCTTCAGTATTGAACCTGGCATCTGCCCTCACTGTGGTGA 2277
Db 1934 CTCCCATACGCGCTCCAGCGGCTTCACGATCGAGCTGGCCCTCAGGCTTCACAGTGGTGA 1993
Qy 2278 TTGCATCAATATTGGCTTCCCATCAGTACACACATTTGAAGTGGGCTCTGTTGTGT 2337
Db 1994 TCGCTTCAACATCGGGCTTCAGTACAGCAGCAGCTGTAAAGTGGGCTCGGTGGTGG 2053
Qy 2338 CTGTTGGCTGGCTCCGGTCCAGGAAGGCTTTGACTGGGCTCTCTTTGTAACATTTTA 2397
Db 2054 CCGTGGGCTGGATCCGCTCCCGAAGGCTGTGACTGGGCTCTTTTCGGAACATCTCG 2113
Qy 2398 TGGCTGGTGTTCACAGTCCCATTTCTGGAGTATCAGTGTGCGCATCGGCAATCT 2457
Db 2114 TGGCTGGTGTGAGACCGTCCCTGTGGCTGGGCTGTTCAGCGCTGTCTCATGGCTCTTC 2173
Qy 2458 TCAGATATGTATCCT 2473
Db 2174 TCATGTATGGATCCT 2189

RESULT 6

US-08-582-719-1
; Sequence 1, Application US/08582719
; Patent No. 5633348
; GENERAL INFORMATION:
; APPLICANT: Johann Dr., Stephen V.
; APPLICANT: Van Zeijl Dr., Marja
; APPLICANT: O'Hara Dr., Bryan M.
; TITLE OF INVENTION: Amphotropic Virus Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: 1937 West Main Street
; CITY: Stamford
; STATE: CT
; COUNTRY: United States of America
; ZIP: 06904-0060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/582,719
; FILING DATE: 04-JAN-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,684
; FILING DATE: 16-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 31937-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-321-2361
; TELEFAX: 203-321-2971
; TELE: 710-474-4059
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 244..2202

US-08-582-719-1
Query Match 18.8%; Score 617.2; DB 1; Length 3175;
Best Local Similarity 60.3%; Pred. No. 2.6e-163;
Matches 1203; Conservative 0; Mismatches 718; Indels 75; Gaps 8;
Qy 496 TGGTGGACTACCTATGGATGCTCATCTGGGCTTCATTTATTCATTTGCTTTGGCATTTCT 555
Db 251 TGGATGAGTATTGTTGGATGGTCAATTTGGGTTTCATCATAGCTTTCACTTTGGCCCTTT 310
Qy 556 CCGTGGAGCCAATGATGTAGCAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGA 615
Db 311 CTGTTGGTGAACACGATGTTGCCAACTCTCTTTGGTACAGCCGTGGGCTCTGGTGTGGTGA 370
Qy 616 CCTGAAGCAAGCCCTGCATCTTAGCTAGCATCTTTGAAACAGTGGGCTCTGCTTACTCG 675
Db 371 CTTGAGGAGGAGCATGCAATTTTAGCTTCAATATTTGAAACACACCGGCTCCGTTGTTACTAG 430
Qy 676 GGGCCAAAGTGAAGCAAAACCATCCGGAAGGGCTTGAATTGACGTGGAGATGTACAACTCGA 735
Db 431 GGGCCAAAGTGAAGCAAAACCATTCGAAAGGTATCATTGACGTGAACCTGTACACAGAGA 490
Qy 736 CTCGAAGGGCTACTGATGGCCGGCTCAGTCAGTGTCTATGTTGGTTCCTGCTGTGGCAAC 795
Db 491 CGGTGGAGACTCTCATGGCTGGGGAAGTTAGTGCCATGGTTGGTTCCTGGCTGTGGCAGC 550
Qy 796 TCGTGGCTTCTGTTTGAAGCTCCCTATTTCTGGAACCATTTGTTGTTGGTGGCACTA 855
Db 551 TGAATGCTTCTCTCTGAGGCTTCCAACTCAGGAACGCACTGCATTTGGGTTCTACTA 610
Qy 856 TTGGTTTCTCCCTCGTGGCAAAAGGGGAGGGGTGTCAAGTGGTCTGAACTGATAAAAA 915
Db 611 TAGGATTCCTCACTGGTGGCAATCGGTACCAAGGTGTGAGTGGAGGCTGTCAAGA 670
Qy 916 TTGTGATGCTTTGGTGTGGTGTCCCACTGCTTTCTGGAATTAATGCTGTGAAATTTATTCT 975
Db 671 TTGTGCTTCTTGGTTTATATCTCCACTGTTCTGGTTTCTATGCTGGCTGCTGCTTTG 730
Qy 976 TCTGGTTCGTGCTATTCCTCCATAGGAGAGATCCAGTTCCTATGCTTTGCGAGCTT 1035
Db 731 TACTCATGAAATTTTCATCTTAAAGGAAGAGACCTGTTCCCAATGGCTCCGGGAC 790
Qy 1036 TGCCAGTTTTCTATGCTGCTGCAAGTTGGAATAAACTCTTTTCCATCATGTATCTGGAG 1095
Db 791 TCCAGTATTCTATGCTGCTACCATAGCAATCAATGCTCTTTTCCATCATGTACACAGGAG 850
Qy 1096 CACCGTTGCTGGGCTTTGACAACTTCCTCTGTGGGTTACCATCTCTCATCTCGTGGGAT 1155
Db 851 CACCAAGTCTCGGCCCTTG---TCTCCCATGTGGGCCCATAGCCCTCATTTCTTTGGTG 907
Qy 1156 GTGCAAGTTTTCTGTCCTTATCGTCTGGTCTTTGATGTCCTCCAGGATGAAGAGAAAA 1215
Db 908 TCGCCCTCTGTTGCTTTTGTGGTCTCTTGTGTTGCTCTTGTGTTGCTGATGCGGAGAAAA 967
Qy 1216 TTGAACGAGAAATAAAGTGTAGTCTCTTGAAGGCCCTTAAATGGAAGAAAAAGATAGCT 1275
Db 968 TAACAGGCAAAATTACAAAAGAGAGGTCTTTATCAGAGTATCTGACGAAGAGCTCAGTA 1027
Qy 1276 TGAAGAAGACCATGAAGAAACAAAGTTGTCTGTTGGTGATATTGAAACCAAGCATCTCTG 1335
Db 1028 AGGTTTCAAGAACAGAGTCCCCAGTATT-----TAAAGAGCTACACAG 1070
Qy 1336 TTTCTGAGGTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1395
Db 1071 TCCCAAGGCTAATGATGACAGCACCATCTCCGCTCAGGAGGAGCAGCGGGAG----- 1122
Qy 1396 CATTCAAACTTGGAGATTTGGAGAGAGCTCCAGAGAGAGAGAGGCTTCCAGCGTGGACT 1455
Db 1123 -----ACACTGGGAGCTCGGAAGGACATCTCTGCG-----GGCAGCCACCTCGGGCT 1170
Qy 1456 TGAAGAGGAAACACAGCATAGATAGCACCGTGAATGCTGAGTGCAGTGTGCTTATGGGA 1515
Db 1171 GCATAGGGAAGAGCACTGTCTCATGCCCATGGCTCTGTGAAATTCGCCCATCTTCCACGGC 1230

QY 1516 ACCTTGTCCAGTTTCAGTCAAGCCGTCAGCAACCAAAATAAACTCCAGTGGCCACTACCAGT 1575
Db 1231 ACCTT-----CGGCTTCAGCGCCACACAGGAGCGACGGTCATGTGT 1273
QY 1576 ATCACACCGTGCATAAGGATTCGGGCTGTACAAAGAGCTACTCCATAAATTACATCTTG 1635
Db 1274 ACCACACCGTGCATAAGACTCGGGGCTCTACAAAGATCTGTGCAAAATCCACATCG 1333
QY 1636 CCAAGTGGGAGATTGCATGGGAGACTCCGGTGACAA-----ACCCTTAAGCGCAATA 1689
Db 1334 ACAGGGGCCCGAGGAGAGCCAGCCAGGAAGCAACTACCGGCTGCTCCGCGGAAACA 1393
QY 1690 ATAGCTATCTCTATACATGGCAATATGTGGCATGCTCTGGATTCATTCCTGGGCCA 1749
Db 1394 ACAGTTTACACTCTACCGCAGCCATTTGTGGCTGCGAGTGCACGCCACCTTTCGAG 1453
QY 1750 AAGAAGTGAACAGAGGGCGGAAGAAATGGAGAAGCTGACATGCCCTAATG---CAGACT 1806
Db 1454 CTGGGACTCATCGGCCCCAGAGGACAGTGAAGAGCTGTGGCGACACCGTGTCTCTACT 1513
QY 1807 CCAAGAAGCGAATTCGAATGGACAGTTACACCAAGTTACTGCAATGTGTCTGACCTTC 1866
Db 1514 CCAAGAAGAGGCTGCGCTACGACAGCTACTCGAGCTACTGTAAACGGGTGGCAGAGCGG 1573
QY 1867 ACTCAGCATCTGAGA-----TAGACATAGTGTCAAGGCAGAGATGGGFTTAGGTG 1917
Db 1574 AGATCAGGCGGAGGAGGCGCGTGGAGATGAAGCTGGCGTGGAGCTGGCGACCCCTG 1633
QY 1918 ACAGAAAGAGAGTAAATGCTCTCTAGAAGATGGTATGACAGGATAAGCTGAAGTCT 1977
Db 1634 ACCAGCGGAGAGGACCCCTGCAGAGGAGGAGAGGAGAGGAGAGGAGGAGGAGGAGGTTTC 1693
QY 1978 CTCTCTCTTCCAGTTCTTCAGATCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGGTG 2037
Db 1694 ACCTCTGTTCATTTCTTCAGGTCCTCACCGCCTGTTTCGGGTCTTTGCTTCACGGCG 1753
QY 2038 GCAATGACGTAGCAATTCGAGTGGCCCTCTGGTCTGCTTTATATTTGGTTTATGACACAG 2097
Db 1754 GCAATGACGTAGTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1813
QY 2098 GAGATGTTTCTTCAAGTGGCAACACCAATATGGCTTCTACTCTATGTGTGTGGTA 2157
Db 1814 GCGGGTAAACGCAAGAGCAGCTACACCCGCTGCGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1873
QY 2158 TCTGTCTGCTCTGCTGGGTGGGGAAGAGTATACAGACCATGGGAGAGTCTGA 2217
Db 1874 TCTGCACAGGCTCTGGGTCTGGGGAGAGAGTATCAGACCATGGGAGAGGACCTCA 1933
QY 2218 CACCGATCACGCTCTAGTGGCTTCAGTATGAACCTGCGATCTGCGCTCAGTGTGGTA 2277
Db 1934 CTCCCATCAGCGGCTCCAGCGGCTTCAGATCAGCTGCGCTCAGGCTTCACAGTGGTGA 1993
QY 2278 TTGCATCAATATGGGCTTCCCATCAGTACACACATGTAAGTGGGCTCTGTGTGT 2337
Db 1994 TCGCCTCCAACTCGGGCTTCCAGTCAAGCAGCAGCTGTAAGTGGGCTCGGTGGTGG 2053
QY 2338 CTGTGGCTGGCTCCGGTCCAGAGGCTGTGACCTGGGCTCTCTTTCTGTAACATTTTA 2397
Db 2054 CCGTGGGCTGGATCCGCTCCGCAAGGCTGTGACCTGGGCGCTCTTTTCGGAACATCTTCG 2113
QY 2398 TGGCTGGTGTGTACAGTCCCATTTCTGGAGTTATCAGTGTGCGCATCATGGCAATCT 2457
Db 2114 TGGCTGGTGTGTGACGCTCCGCAAGGCTGTGACCTGGGCGCTCTTTTCGGAACATCTTCG 2173
QY 2458 TCGATATGATCTCT 2473
Db 2174 TCATGTATGGATCCT 2189

Patent No: 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 623
LENGTH: 662
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(662)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-623

Query Match 15.1%; Score 495.2; DB 4; Length 662;
Best Local Similarity 90.2%; Pred. No. 2.1e-129;
Matches 591; Conservative 0; Mismatches 54; Indels 10; Gaps 6;
QY 1606 ACAAGAGCTACTCCATAAAATTACATCTTGCCAAAGTGGGAGATTGTCATGGGAGACTCCG 1665
Db 1 ACANAGAGCTACTCCATAAAATTACATCTTGCCAAAGTGGGAGATTGTCATGGGAGACTCCG 60
QY 1666 GTGCAACACCCCTTAAGGCGCAATAATAGCTATATCTCTATACCATGGCAATATGCGCA 1725
Db 61 GTGCAACACCCCTTAAGGCGCAATAATAGCTATATCTCTATACCATGGCAATATGCGCA 120
QY 1726 TCCCTCTGATTCATTCCGTGCCAAAGAGGTGAACAGAGCGCGGAAGAAATGGGAAGC 1785
Db 121 TCCCTCTGATTCATTCCGTGCCAAAGAGGTGAACAGAGCGCGGAAGAAATGGGAAGC 180
QY 1786 TGACATGGCCTAATGCAGACTCCAAAGAGCGAATTCGAATGGACAGTTACACAGTTACT 1845
Db 181 TGACATGGCCTAATGGGACTCCNAGAGCGAATTCGAATGGACAGTTACACAGTTACT 240
QY 1846 GCAATGCTGTGTCTGACCTTCACTCAGCATCTGAGATGACATGATGTCAAGGCGAGAGA 1905
Db 241 GCAATGCTGTGTCTGACCTTCACTCAGCATCTGAGATGACATGATGTCAAGGCGAGAGA 300
QY 1906 TGGGCTAGGTGACAGAAAGGAA--GTAATGGCTCTCTAGAAAGTGGTATGACCAGGA 1963
Db 301 TGGGCTAGGTGACAGAAAGGAAAGTAATGGGCTCTCTAGAAAGTGGTATGACCAGGA 360
QY 1964 TAAGCTGAAGTCTCTCTCTCTCCAGTTCCTGAGATCCCTTACAGCCTGCTTTGGGTC 2023
Db 361 TAAGCTGAAGTCT 420
QY 2024 ATTCGCCCATGGTGGCAATGACGTAAAGCAATGCCA--TTGGGCTCTGGTGTCTTTATTT 2082
Db 421 ATTCGCCCATGGTGGCAATGACGTAAAGCCATGCCATTTGGGCTCTGGTGTCTTTATTT 480
QY 2083 TGGTTTATGACACAGAGATGTTTCTTCAAAAGTGGCAACCAATA--TGGCTTCTACTTC 2141
Db 481 TGGGTTATGACCCNNGAGANG--TTCTTCAAAAGTGGCAACCAATAATTTGGTCTTCTACTTC 539

QY 2142 TAT--GGTGTGTGTGTATCTGTGTGG---TCTGTGGGTTTGGGGAAGAAGAGTTATCC 2196
|||
Db 540 TANTGNGGGGTTGGATCTGNGGTTGGTCTGTGGGTTTGGGGAAGAAAAGTTTCC 599
|||
QY 2197 AGACCATGGGAAGGATCTGACACCCATCACCCCTCTAGTGGCTTCAGTATTGA 2251
|||
Db 600 CNACTTGGGGAAGGATTTGCCNCGTTACACCCCTTTAGGGGTTTNGTATTGA 654
|||

RESULT 8

US-09-328-111-521
; Sequence 521, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 521
; LENGTH: 613
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(613)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-521

Query Match 13.7%; Score 451.4; DB 4; Length 613;
Best Local Similarity 99.2%; Pred. No. 4.2e-117;
Matches 474; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1131 GGTACCATCTCATCTCGGTGGGATGTCAGTTTCTGTGCCCTTATCGTCTGGTCTTT 1190
Db 1 GGTACCATCTCATCTCGGTGGGATGTCAGTTTCTGTGCCCTTATCGTCTGGTCTTT 60
|||
QY 1191 GTATGTCCCAGGATGAAGAGAAAATTGAACGAGAAATAAGTGTAGTCTCTGAAAGC 1250
Db 61 GTATGTCCCAGGATGAAGAGAAAATTGAACGAGAAATAAGTGTAGTCTCTGAAAGC 120
|||
QY 1251 CCCTTAATGAAAAAAGATAGCTTGAAGAAGACCATGAAGAAACAAAGTTGTCTGTT 1310
Db 121 CCCTTAATGAAAAAAGATAGCTTGAAGAAGACCATGAAGAAACAAAGTTGTCTGTT 180
|||
QY 1311 GGTGATATTCAAACAGCATCTGTTTCTGAGTAGGGCTCGCATCTGCGCCCTCCAG 1370
Db 181 GGTGATATTCAAACAGCATCTGTTTCTGAGTAGGGCTCGCATCTGCGCCCTCCAG 240
|||
QY 1371 GCTGTGGTGAGAGAGAACAGTCTCAATCAAACTTGGAGATTGGAGGAGCTCCAGAG 1430
Db 241 GCTGTGGTGAGAGAGAACAGTCTCAATCAAACTTGGAGATTGGAGGAGCTCCAGAG 300
|||
QY 1431 AGAGAGAGCTTCCAGCGTGGACTTGAAGAGAGAACAGCATAGATACACCGTGAAT 1490
Db 301 AGAGAGAGCTTCCAGCGTGGACTTGAAGAGAGAACAGCATAGATACACCGTGAAT 1490
|||

QY 1491 GGTGAGTGCAGTTCCTTAATGGGAACCTTGTCCAGTTTCAGTC-AAGCCGTGAGAACCA 1549
|||
Db 361 GGTGAGTGCAGTTCCTTAATGGGAACCTTGTCCAGTTTCAGTC-CAAGCCGTGAGAACCA 420
|||
QY 1550 AATAAACTCCAGTGGCCACTACAGTATCACCCGTGCAT-AAGGATTCGGGCTGTGA 1606
|||
Db 421 AATAAACTNCACTGGCCACTACAGTATCACCCGTGCAT-AAGGATTCGGGCTGTGA 478
|||

RESULT 9

US-09-328-111-269/c
; Sequence 269, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 269
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(643)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-269

Query Match 11.9%; Score 393; DB 4; Length 643;
Best Local Similarity 91.7%; Pred. No. 1.2e-100;
Matches 465; Conservative 0; Mismatches 35; Indels 7; Gaps 5;
QY 1805 CTCCAAGAAGCGAATTCGAATGGACAGATTACACCACTTACTGCAATGCTGTCTGACCT 1864
Db 505 CTCCAAGAAGCGAATTCGAATGGACAGATTACACCACTTACTGCAATGCTGTCTGACCT 448
|||
QY 1865 TCACCTCAGCATCTG-AGATAGACATGAGTGTCAAGCAGAGATGGGTCTAGGTGACAGAA 1923
Db 447 TCATTNAGCNVTTTAAGTAGCCATAAGTGCAAGNCAAGATGGGTTTAGGTGNCAGAA 388
|||
QY 1924 AAGGAAGTAATGGCTCT--CTAGNAGAATGGTATGACCAAGGATAG-CCTGAAGTCTCTC 1980
Db 387 AAGAAAGTAAGGCTTCTCTAGNAGAATGGTATGACCAAGGATAG-CCTGAAGTCTCTC 328
|||
QY 1981 TCC-TCCTTCAGTTCCTGCAGATCCTTACAGCTGCTTTGGGTGCTATTCCGCCATGGTGC 2039
Db 327 TCCTTCCTTCAGTTCCTGCAGATCCTTACAGCTGCTTTGGGTGCTATTCCGCCATGGTGC 268
|||
QY 2040 AATGACGTAAGCAATGCCATTGGGCTCTGGTGTCTTTATATTTGGTTTATGACAGGA 2099
Db 267 AATGACGTAAGCAATGCCATTGGGCTCTGGTGTCTTTATATTTGGTTTATGACAGGA 208
|||
QY 2100 GATGTTTCTTCAAAAGTGGCAACCAATATGGCTCTACTCTATGCTGTGTGTATC 2159
Db 207 GATGTTTCTTCAAAAGTGGCAACCAATATGGCTCTACTCTATGCTGTGTGTATC 148
|||

QY 2160 TGTGTTGGTCTGTGGTGTGGGGAAGAGTTATCCAGACCATGGGGAAGGATCTGACA 2219
Db 147 TGTGTTGGTCTGTGGTGTGGGGAAGAGTTATCCAGACCATGGGGAAGGATCTGACA 88
QY 2220 CCATCACACCCCTCTAGTGGCTTCAGTATTGAACCTGGCATCTGCCCTCACTGGTGATT 2279
Db 87 CCATCACACCCCTCTAGTGGCTTCAGTATTGAACCTGGCATCTGCCCTCACTGGTGATT 28
QY 2280 GCATCAAAATATTGGCCTTCCCATCAGT 2306
Db 27 GCATCAAAATATTGGCCTTCCCATCAGT 1

RESULT 10
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 2.2%; Score 73.6; DB 4; Length 4403765;
Best Local Similarity 50.3%; Pred. No. 1.9e-08;
Matches 181; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 513 ATGCTCATCTGGGCTTCATTATTCGATTTGTCTGGCATTTCTCGTGGGAGGCAATGAT 572
Db 2549186 ATCTGGTGTCTGCCACCGCTGTCGGCATGTTTCATGGCCTTCAACGTCGGCGGCATGAT 2549245

QY 573 GTAGCAAAATCTTTGGTACAGCTGGGCTCAGGTGTAGTGACCTGAAGCAAGCCTGC 632
Db 2549246 GTCCCAACTCGTTTGGCACCCAGCGTCGGCGGGGACGTTGACCATGAACAGGCGCTT 2549305

QY 633 ATCTAGTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCCAAAGTGAGCGAA 692
Db 2549306 CTGGTCGGCGGATCTTCGAGGTACGGGGCGGCTATCGCCGGCGGACGTCACCGAG 2549365

QY 693 ACCATCCGAAGGGCTTGATTGACGTGGAGATGACAACTCGACTCAAGGGCTACTGATG 752
Db 2549366 ACCATCCGAGCGCATCGTTGATCTGTCCGGGCTGTCCGTGACCCACGCGACTTTCATG 2549425

QY 753 GCCGGCTCAGTCAGTCTATGTTGGTTCTGTGTGGCAACTCGTGGCTTCGTTTTTG 812
Db 2549426 AACATCATGCTGTCGGCGCTATCGGCAGCGCGCTCTGGCTGTGTTTGTAAACCGTATG 2549485

QY 813 AAGCTCCCTATTCTGGAACCCATTGATTGTTGGTGCAACTATTGTTTCTCCCTCGTG 872
Db 2549486 GGGTACCCGGTGTGACACACACTGATCATCGGCGGCATGTCGGCGGCGATCGCG 2549545

RESULT 11
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 2.2%; Score 73.6; DB 4; Length 4411529;
Best Local Similarity 50.3%; Pred. No. 1.9e-08;
Matches 181; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 513 ATGCTCATCTGGGCTTCATTATTCGATTTGTCTGGCATTTCTCGTGGGAGGCAATGAT 572
Db 2553375 ATCTGGTGTCTGCCACCGCTGTCGGCATGTTTCATGGCCTTCAACGTCGGCGGCATGAT 2553434

QY 573 GTAGCAAAATCTTTGGTACAGCTGGGCTCAGGTGTAGTGACCTGAAGCAAGCCTGC 632
Db 2553435 GTCCCAACTCGTTTGGCACCCAGCGTCGGCGGGGACGTTGACCATGAACAGGCGCTT 2553494

QY 633 ATCTAGTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCCAAAGTGAGCGAA 692
Db 2553495 CTGGTCGGCGCATCTTCGAGGTACGCGCGGGTGTATCGCCGGGCTGTCGTCGCGCGGATCGG 2553554

QY 693 ACCATCCGAAGGGCTTGATTGACGTGGAGATGACAACTCGACTCAAGGGCTACTGATG 752
Db 2553555 ACCATCCGAGCGCATCGTTGATCTGTCCGGGCTGTCGTCGACCCACGCGACTTTCATG 2553614

QY 753 GCCGGCTCAGTCAGTCTATGTTGGTTCTGTGTGGCAACTCGTGGCTTCGTTTTTG 812
Db 2553615 AACATCATGCTGTGGCGCTATCGGCAGCGCGGCTCTGGCTGTGTTGTAACCGTATG 2553674

QY 813 AAGCTCCCTATTCTTGAAACCCATTGATTGTTGGTGCAACTATTGTTTCTCCCTCGTG 872
Db 2553675 GGGTACCCGGTGTGACACACACTGATCATCGGCGGCATGTCGCGCGGCGATCGCG 2553734

RESULT 12
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22133-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: EP 91 114 300.6
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match 1.9%; Score 62.2; DB 1; Length 7218;
Best Local Similarity 5.3%; Pred. No. 6e-07;
Matches 22; Conservative 230; Mismatches 163; Indels 0; Gaps 0;

QY 18 TTCTTCTTCGCGCGCTGAACCCCGCGGCTCTCTCGGGAAGTCTGAGTCCCGC 77
Db YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1128
QY 78 TGAGCTGTCCCGCTGCGCGCGAGCCGCGGTGCGCGTGTCTCGAGCGCTGCGGC 137
Db YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1188
QY 138 TCGATCTCTGCTCCCGCTCGCGCTCCCTTTTCCCTGGATGAATGCGCTTCT 197
Db YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1248
QY 198 CTCTCGCGCATGGAATCTGCTCGGCTTTTAGCCCTCTGAGCCCTCTGAGCAAGAACCCAG 257
Db YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1308
QY 258 ACAACAGATGCCATAGCAGCGGTATAGCAGTAAGTAACTCCCGAGCTCGTGTCTGCGGTA 317
Db YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1368
QY 318 GTTTACAGTATTTAATTTATATAATATATATTTATTTATTTATAGCATTTTGTATACCTCA 377
Db YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1428
QY 378 TATCTGTTTACATCTTGAAGGCGCTCAGTAGTCTCTTACTAAACACAC 432
Db YYYYYYGTACCAATTTCTCTATCTCTTTAACTACTTGCATAGTAGGTAATTAC 1483

RESULT 13
US-09-134-001C-2591
; Sequence 2591, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2591

; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-2591

Query Match 1.8%; Score 58.6; DB 4; Length 1029;
Best Local Similarity 47.2%; Pred. No. 1.9e-06;
Matches 212; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

QY 1986 TTCCAGTTCTTCGAGATCCTTACAGCCTCTTTGGGTGCTATTCGCCCATGGTGGCAATGAC 2045
Db TTTAGGTTTTTCCAAATCTTTACTGCAGCATTAACATCAATCTCACATGGTACAACAGAT 600
QY 2046 GTAAGCAATGCCATTGGGGCTCTGGTGTCTTTATATTTGGTTTATGACACAGAGATGTT 2105
Db GCTCAAAATCTATGGG---TATTATCACATTAGCTTTAATTGTAGGTAATTTACAAGAT 657
QY 2106 TCTTCAAAAGTGGCAACACCAATATGGTTTCTACTCTATGTTGGTGTGTTGTTGTTGTT 2165
Db GGTAGTATGTTGAACCTCAAGTATGGGTTAAAGTTGCGTGTGCAACAGCTATGGGATTA 717
QY 2166 GGTCTGTGGGTTTGGGGAAGAGATTATCCAGACCATGGGGAAGGATCTGACACCGATC 2225
Db GGTACAGCAGTAGGTGGTTGGAAATCATCAAAATCTGTGGCGGAAATATATGAAAT 777
QY 2226 ACACCTCTAGTGGCTTCAGTATTGAACCTGGCATCTGCCCTCACTGTGGTGTGATGTCATCA 2285
Db AGACCTCTTAACGGTGCAGCCGCTGATATTTCTTCAGCATTAACGATTTTCGTAGGCTCA 837
QY 2286 AATATTGGCTTCCCATCAGTACACACATTTGTAAGTGGGCTCTGTGTGTCTGTGTGTC 2345
Db TCATTACATTTCCCATTTACTACTACACACAGTGTGTCTTCTCATCTATTCTAGGTGTAGT 897
QY 2346 TGCTCCCGTCCAAAGGCTGTTGACTGGCGTCTTTTCGTAACATTTTATGCGCTCG 2405
Db GTTTCAAATCGTGCGAAAGGTGTTAAATGGAGTACTGCTCAACGATGTTGTGTAACATGG 957
QY 2406 TTTGTCACAGTCCCATTTTCTGGAGTTAT 2434
Db GTGATTACTTTACCTATCTCAGCTGTTTT 986

RESULT 14
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
```

```

; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: BENT, Stephen A.
; ; REGISTRATION NUMBER: 29,768
; ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (703)836-9300
; ; TELEFAX: (703)683-4109
; ; TELEX: 899149
; ; INFORMATION FOR SEQ ID NO: 14:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 7218 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; IMMEDIATE SOURCE:
; ; CLONE: ptzgmt-F18
; ; US-08-232-463-14

```

Query Match		1.6%	Score 52.2;	DB 1;	Length 7218;	
Best Local Similarity		4.6%	Pred. No. 0.00039;			
Matches	18;	Conservative	214;	Mismatches	157;	Indels 0; Gaps 0;

Qy	1200	AGGATGAAGAAGAAAAATTGAAACAGAGAAATAAAGTGTAGTCCTTGAAAAGCCCTTAATG	1259
Dy	1420	RR	1361
Qy	1260	GAAAAAAGAATAGCTTTGAAAGAAGACCATTGAAGAAAACAAAGTTGCTGTGGTGATATT	1319
Dy	1360	RR	1301
Qy	1320	GAAAAACAAGCATCCTGTTTTCTGAGTAGGCCCTGCCACTTGCCCCCTCCAGGCTGTGGTG	1379
Dy	1300	RR	1241
Qy	1380	GAGGAGAGAACAGTCTCATTTCAAACTTTGAGATTGAGAGAAGCTCCAGAGAGAGAGAGG	1439
Dy	1240	RR	1181
Qy	1440	CTTCCCAGCGTGCACTTTGAAAGAGAAACACAGATAGATAGCACCGTGAATGGTGCAGTG	1499
Dy	1180	RR	1121
Qy	1500	CAGTGCCTAATGGGAACCTTGTCAGTTCAGTCAAGCCGTGAGCAACCAAAATAAAGCTCC	1559
Dy	1120	RR	1061
Qy	1560	AGTGGGCATTCACGATATCACCGTGCA	1588
Dy	1060	AGCTCCCTTCAGCTGCAGCCAAGCTCGGA	1032

```

RESULT 15
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match      1.3%; Score 42.2; DB 4; Length 289;
Best Local Similarity 6.0%; Pred. No. 0.035;
Matches 14; Conservative 106; Mismatches 113; Indels 0; Gaps 0

Qy      4  CGCTCTCTGGTGGTTCCTCTCTCGCGCGCTGAAACCCCGCGGCTGCTTCTCTGGGAAG 63
Db      233  YGYCTAYGCVTYGYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYN 174

Qy      64  GTCGTGAGTCCCGCTGAGCTGTCCCCGGTGCGCGGACCCCGGCGCTGTGCGCGTGGCTC 123
Db      173  YNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYN 114

Qy      124  CAGCGCGTGGCGCCTCGATCTCTGCTCTCCCGCTCCGCGCCTCCCTTTTCCCTGATGAA 183
Db      113  YNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYN 154

Qy      184  CTTGCGTCCCTTCTCTCTCCGCGCATGGAATCTGCTCCGTCCTTTAGCCCT 236
Db      53  YTTGTGYTAYATYTTGYTAYAYATYAYATYAYGYTAYATYTGTYTCYCYC 1

```

Search completed: June 28, 2003, 08:20:19
Job time : 133.651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 06:24:51 ; Search time 310.679 Seconds
(without alignments)
15719.871 Million cell updates/sec

Title: US-09-981-353-172
Perfect score: 3290
Sequence: 1 cggcgctctctcggtgttc.....agtttttagaaaaa 3290

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/1/pubpna/PCR_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/1/pubpna/PCRUS_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3290	100.0	3290	9	US-09-981-353-172 Sequence 172, App
2	3290	100.0	3290	9	US-09-919-039-164 Sequence 164, App
3	3175.2	96.5	3220	9	US-09-954-531-137 Sequence 137, App
4	3175.2	96.5	3220	9	US-09-954-531-137 Sequence 356, App
5	495.2	15.1	662	10	US-09-879-536-623 Sequence 623, App
6	459.8	14.0	463	10	US-09-867-701-2803 Sequence 2803, App
7	451.4	13.7	613	10	US-09-879-536-521 Sequence 521, App
8	448	13.6	495	12	US-09-920-300A-1304 Sequence 1304, App
9	448	13.6	495	12	US-10-033-528-1304 Sequence 1304, App
10	430	13.1	441	10	US-09-867-701-2808 Sequence 2808, App
11	406.4	12.4	414	10	US-09-867-701-3261 Sequence 3261, App
12	394.4	12.0	396	10	US-09-867-701-2909 Sequence 2909, App
13	394.2	12.0	404	9	US-09-918-995-17414 Sequence 17414, A
14	393	11.9	643	10	US-09-879-536-269 Sequence 269, App
15	324.4	9.9	338	10	US-09-867-701-9834 Sequence 9834, App
16	306.4	9.3	330	10	US-09-867-701-3723 Sequence 3723, App
17	228.4	6.9	230	9	US-10-102-524-616 Sequence 616, App
18	112.6	3.4	393	9	US-10-198-846-7182 Sequence 7182, App
19	102.8	3.1	1764	9	US-09-938-842A-934 Sequence 934, App

ALIGNMENTS

RESULT 1

US-09-981-353-172
; Sequence 172, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 172
; LENGTH: 3290
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 2023119CBI
US-09-981-353-172

Query Match	100.0%	Score	3290;	DB	9;	Length	3290;
Best Local Similarity	100.0%;	Pred. No. 0;					
Matches	3290;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	CGCGCGCTCTCTCGGTGGTTCCTCTCGCGCGCTGAAACCCCGCGGTCTCTCTGGG	60				
Db	1	CGCGCGCTCTCTCGGTGGTTCCTCTCTCGCGCGCTGAAACCCCGCGGTCTCTCTGGG	60				
QY	61	AAGGTGCTGAGTCCCGCTGAGTGTCTCCCGGTGCGCCGACCCGCGCGGTGTGTCCCGTGG	120				
Db	61	AAGGTGCTGAGTCCCGCTGAGTGTCTCCCGGTGCGCCGACCCGCGCGGTGTGTCCCGTGG	120				
QY	121	CTCCAGCGCGCTCGCGCTCGATCTCTCTCTCCCGCTCGCGCTCCCTTTTCCCTGGAT	180				
Db	121	CTCCAGCGCGCTCGCGCTCGATCTCTCTCTCCCGCTCGCGCTCCCTTTTCCCTGGAT	180				
QY	181	GAACTTGGCTCTCTCTCTCTCTCCCGCATGGAATCTGCTCGGTGCTTTTAGCCCTCCCTG	240				
Db	181	GAACTTGGCTCTCTCTCTCTCTCCCGCATGGAATCTGCTCGGTGCTTTTAGCCCTCCCTG	240				


```
Db 2401 CCTGGTTTGTACAGTCCCAATTTCTGGAGTTATCAGTGTGCGCATCATGGCAATCTTCA 2460
Qy 2461 GATATGTCATCTCTCAGAACTGAAGCTGTTTGAATTAATAATTTGTCAATTTTGGGA 2520
Db 2461 GATATGTCATCTCTCAGAACTGAAGCTGTTTGAATTAATAATTTGTCAATTTTGGGA 2520
Qy 2521 CCATCTAGGTATCTCTGCTCCCTCAGAACTGTTTACAGTGTATACAGAACTGACAA 2580
Db 2521 CCATCTAGGTATCTCTGCTCCCTCAGAACTGTTTACAGTGTATACAGAACTGACAA 2580
Qy 2581 GAGTCTTTTATTTGGGAGCCAGAGGAGGAGTGTACTTGTGCTATAACTCTTTTGT 2640
Db 2581 GAGTCTTTTATTTGGGAGCCAGAGGAGGAGTGTACTTGTGCTATAACTCTTTTGT 2640
Qy 2641 GCTAAATATGAATGCTCAAAATAGCTGTGTAAATAGCCGGGTTCACCTGGCTCCT 2700
Db 2641 GCTAAATATGAATGCTCAAAATAGCTGTGTAAATAGCCGGGTTCACCTGGCTCCT 2700
Qy 2701 GCTGAGGTCCCTTTCTCTGCGCTGTGAATTCCTGTACATATTTCTCTACTTTTGT 2760
Db 2701 GCTGAGGTCCCTTTCTCTGCGCTGTGAATTCCTGTACATATTTCTCTACTTTTGT 2760
Qy 2761 TCAGGCTTCAATTCATATGTTTAAATGTTGCTCTGAAGATGACTTGTGATTTTGT 2820
Db 2761 TCAGGCTTCAATTCATATGTTTAAATGTTGCTCTGAAGATGACTTGTGATTTTGT 2820
Qy 2821 TCTTTTTTAAACCATAGAGCCGTTTGACAGACATGCTCTGCGTTTGTGGTTTCAC 2880
Db 2821 TCTTTTTTAAACCATAGAGCCGTTTGACAGACATGCTCTGCGTTTGTGGTTTCAC 2880
Qy 2881 CAGCTTCTGCGCTCAGATGACAGGAGTTTAAACAAATAATATACAACTTCCCTT 2940
Db 2881 CAGCTTCTGCGCTCAGATGACAGGAGTTTAAACAAATAATATACAACTTCCCTT 2940
Qy 2941 GTAGTCTCTTATATAAGTAGAGTCTTGTGACTCTGCGCTCTGCTCAGTAGTGGCAGGAT 3000
Db 2941 GTAGTCTCTTATATAAGTAGAGTCTTGTGACTCTGCGCTCTGCTCAGTAGTGGCAGGAT 3000
Qy 3001 CTATTGGCATATTCGGGAGCTCTTAGAGGATGAGGTTCTTTGAACACAGTGAATAATTT 3060
Db 3001 CTATTGGCATATTCGGGAGCTCTTAGAGGATGAGGTTCTTTGAACACAGTGAATAATTT 3060
Qy 3061 AAATTAGTAATTTTTCAGAGCTTTTATGACTGTTTATGCTTAAGAGAGTAAGAAA 3120
Db 3061 AAATTAGTAATTTTTCAGAGCTTTTATGACTGTTTATGCTTAAGAGAGTAAGAAA 3120
Qy 3121 GAAAAAGCTGTTGGCAATCTCTGTTATTTCTTTAAGATTTCTGGCAGTGTGGGAT 3180
Db 3121 GAAAAAGCTGTTGGCAATCTCTGTTATTTCTTTAAGATTTCTGGCAGTGTGGGAT 3180
Qy 3181 GAATGAAGTGAATGTGAATTTGGCAAGTTAAATGGGACAGCTTCCATGTTCAATTG 3240
Db 3181 GAATGAAGTGAATGTGAATTTGGCAAGTTAAATGGGACAGCTTCCATGTTCAATTG 3240
Qy 3241 TCTACCTTAACTGAATAAAAAAGCTCAGTTTTTTAGAAAAA 3290
Db 3241 TCTACCTTAACTGAATAAAAAAGCTCAGTTTTTTAGAAAAA 3290
```

RESULT 2

```
US-09-919-039-164
; Sequence 164, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
```

```
; SEQ ID NO 164
; LENGTH: 3290
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 2023119CBI
US-09-919-039-164
```

```
Query Match 100.0%; Score 3290; DB 9; Length 3290;
Best Local Similarity 100.0%; Pred No. 0;
Matches 3290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCTCTCTCGTGGTTCCTTCTCGGCCGCTGAAACCCCGCGGTGCTTCTCTGGG 60
Db 1 CGCGCGCTCTCTCGTGGTTCCTTCTCGGCCGCTGAAACCCCGCGGTGCTTCTCTGGG 60
Qy 61 AAGTCTGTGAGTCCCGCTGAGCTGTCCCGGTGCCCGCAGCCGGGCGGTGTCCTGGTGG 120
Db 61 AAGTCTGTGAGTCCCGCTGAGCTGTCCCGGTGCCCGCAGCCGGGCGGTGTCCTGGTGG 120
Qy 121 CTCCAGCCGCTCGCGCTCGATCTCTCTGCTCCCGCTCCGCCCTCCCTTTTCCCTGGAT 180
Db 121 CTCCAGCCGCTCGCGCTCGATCTCTCTGCTCCCGCTCCGCCCTCCCTTTTCCCTGGAT 180
Qy 181 GAATCTGTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 181 GAATCTGTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy 241 AGCCAAAGAAACCCAGACAGATGCCCATACGACAGCTATAGCAGTAATCTCCCGCAGC 300
Db 241 AGCCAAAGAAACCCAGACAGATGCCCATACGACAGCTATAGCAGTAATCTCCCGCAGC 300
Qy 301 TCGGTTTCTGTCGCTAGTTTACAGTATTTTAAATTTTATATAATATATATTTATTTATTA 360
Db 301 TCGGTTTCTGTCGCTAGTTTACAGTATTTTAAATTTTATATAATATATATTTATTTATTA 360
Qy 361 GCATTTTGTATACCTCATATTTCTTTTACACATCTTTGAAAGGCGCTCAGTAGTTCTCTTA 420
Db 361 GCATTTTGTATACCTCATATTTCTTTTACACATCTTTGAAAGGCGCTCAGTAGTTCTCTTA 420
Qy 421 CTAAACACCACTACTCCAGAGATGGCAACGGTGTATACCACTACTACAGCTGTACCG 480
Db 421 CTAAACACCACTACTCCAGAGATGGCAACGGTGTATACCACTACTACAGCTGTACCG 480
Qy 481 CGCTTCTGCTCTTCTGTTGGTACTACCTATGATGCTCATCTCTGGCTTCATTATTGCTAT 540
Db 481 CGCTTCTGCTCTTCTGTTGGTACTACCTATGATGCTCATCTCTGGCTTCATTATTGCTAT 540
Qy 541 TTGCTTGTGCATTTCTCGTGGAGCAATGATGTAGCAAAATTTCTTTTGTACAGCTGTGG 600
Db 541 TTGCTTGTGCATTTCTCGTGGAGCAATGATGTAGCAAAATTTCTTTTGTACAGCTGTGG 600
Qy 601 GCTCAGGTGTAGTACCCCTGAAAGACCTGATCTCTAGCTAGCATCTTTTGAACAGTGG 660
Db 601 GCTCAGGTGTAGTACCCCTGAAAGACCTGATCTCTAGCTAGCATCTTTTGAACAGTGG 660
Qy 661 GCTCTGCTTACTGGGGCCAAAGTGAAGCAACCATCCGGAAGGCTTGTATTGACGTGG 720
Db 661 GCTCTGCTTACTGGGGCCAAAGTGAAGCAACCATCCGGAAGGCTTGTATTGACGTGG 720
Qy 721 AGATGTACAACTCGACTCAAGGGCTACTGATGGCGGCTCAGTCAAGTGTCTATGTTTGGTT 780
Db 721 AGATGTACAACTCGACTCAAGGGCTACTGATGGCGGCTCAGTCAAGTGTCTATGTTTGGTT 780
Qy 781 CTGCTGTGTGGCAACTCTGTTGGTTCCTTTTGAAGCTCCCTATTTCTGGAAACCCATTGA 840
Db 781 CTGCTGTGTGGCAACTCTGTTGGTTCCTTTTGAAGCTCCCTATTTCTGGAAACCCATTGA 840
Qy 841 TTGTTGGTCAACTATTGGTTTCTCCCTGTCGCAAGGGGAGGAGGCTGCAAGTGGT 900
Db 841 TTGTTGGTCAACTATTGGTTTCTCCCTGTCGCAAGGGGAGGAGGCTGCAAGTGGT 900
```

901 CTGAACGATGATAAAATGATGATCTCTGGTTCGTGCTCCACGCTCTTCTGGAATATGT 960
901 CTGAACGATGATAAAATGATGATCTCTGGTTCGTGCTCCACGCTCTTCTGGAATATGT 960
961 CTGGAATTTTATTTCTTCTGCTGGTTCGTGCAATTCATCTCATTAAGGAGAGATCCAGTTCTTA 1020
961 CTGGAATTTTATTTCTTCTGCTGGTTCGTGCAATTCATCTCATTAAGGAGAGATCCAGTTCTTA 1020
1021 ATGGTTTGGAGCTTTGGCAGATTTCTATGCTCTGCACAGTTGGAATAAACCCTCTTTTCCA 1080
1021 ATGGTTTGGAGCTTTGGCAGATTTCTATGCTCTGCACAGTTGGAATAAACCCTCTTTTCCA 1080
1081 TCATGTATACCTGAGCACCCTGCTGCGCTTTGACAAACTTCTCTGCGGGTACCATCC 1140
1081 TCATGTATACCTGAGCACCCTGCTGCGCTTTGACAAACTTCTCTGCGGGTACCATCC 1140
1141 TCATCTCGGTGGAGTGTGCAAGTTTCTGTCGCCCTTATCGTCTGGTTCCTTTGTATGTCCCA 1200
1141 TCATCTCGGTGGAGTGTGCAAGTTTCTGTCGCCCTTATCGTCTGGTTCCTTTGTATGTCCCA 1200
1201 GGATGAAGAGAAATTTGAACGAGAAATTAAGTGTAGTCTCTGAAAGCCCTTAATGG 1260
1201 GGATGAAGAGAAATTTGAACGAGAAATTAAGTGTAGTCTCTGAAAGCCCTTAATGG 1260
1261 AAAAAAGAAATAGCTTTGAAAGAGAGACCATGAAGAAACAAAGTTGTCTGTTGGTGAATTTG 1320
1261 AAAAAAGAAATAGCTTTGAAAGAGAGACCATGAAGAAACAAAGTTGTCTGTTGGTGAATTTG 1320
1321 AAAAAAGCAATCTGTTTCTGAGGTAGGGCTGCCACTGTGCCCCCTCCAGGCTGTGGTGG 1380
1321 AAAAAAGCAATCTGTTTCTGAGGTAGGGCTGCCACTGTGCCCCCTCCAGGCTGTGGTGG 1380
1381 AGGAGAGACAGTCTCAATCAAACTTGGAGATTTGGAGGAAGTCTCCAGAGAGAGAGAGGC 1440
1381 AGGAGAGACAGTCTCAATCAAACTTGGAGATTTGGAGGAAGTCTCCAGAGAGAGAGAGGC 1440
1441 TTCCAGCGTGGACTTTGAAAGAGAGAAACAGCATAGATAGCACCGTGAATGGTGCAGTGC 1500
1441 TTCCAGCGTGGACTTTGAAAGAGAGAAACAGCATAGATAGCACCGTGAATGGTGCAGTGC 1500
1501 AGTTGCTTAATGGGAACCTTGTCCAGTTTCAGTCAAGCCGTTCAGCAACCAAAATAAATCCCA 1560
1501 AGTTGCTTAATGGGAACCTTGTCCAGTTTCAGTCAAGCCGTTCAGCAACCAAAATAAATCCCA 1560
1561 GTGGCCACTACCAAGTATCACACCGTGCATGAAGATTTCCGGCCCTGTCAAAAGAGCTACTCC 1620
1561 GTGGCCACTACCAAGTATCACACCGTGCATGAAGATTTCCGGCCCTGTCAAAAGAGCTACTCC 1620
1621 ATAAATTTACATCTTGCCNAGGTGGAGATTTGCATGGAGACTCCGGTGCACAAACCCCTTAA 1680
1621 ATAAATTTACATCTTGCCNAGGTGGAGATTTGCATGGAGACTCCGGTGCACAAACCCCTTAA 1680
1681 GGGCCAAATAGCTATACCTTCTATACCATGGAATATGTCGATGCGCTCTGGAATTCAT 1740
1681 GGGCCAAATAGCTATACCTTCTATACCATGGAATATGTCGATGCGCTCTGGAATTCAT 1740
1741 TCCGTGCCAAGAGGTGAACAGAGGGCCAGAAATGGAGAGCTGCATGCGCCCTAAATG 1800
1741 TCCGTGCCAAGAGGTGAACAGAGGGCCAGAAATGGAGAGCTGCATGCGCCCTAAATG 1800
1801 CAGACTCCAAAGAGCGAATTCGAATGGACAGTGTACACCAAGTTACTGCAATGCTGTGCTG 1860
1801 CAGACTCCAAAGAGCGAATTCGAATGGACAGTGTACACCAAGTTACTGCAATGCTGTGCTG 1860
1861 ACCTTCACTCAGCATCTGAGATAGACATGAGTGTCAAGGAGAGATGGGTCTAGGTGACA 1920
1861 ACCTTCACTCAGCATCTGAGATAGACATGAGTGTCAAGGAGAGATGGGTCTAGGTGACA 1920
1921 GAAAGGAGTAAATGCTCTAGAGAAATGGTATGACAGGATAGCCCTGAAGTCTCTC 1980
1921 GAAAGGAGTAAATGCTCTAGAGAAATGGTATGACAGGATAGCCCTGAAGTCTCTC 1980
1981 TCCTCTTCCAGTTCTCTGAGATCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGGTGCA 2040

1981 TCCTCTTCCAGTTCTCTGAGATCCTTACAGCCTGCTTTGGGTCAATTCGCCCATGGTGCA 2040
2041 ATGAGCTAAGCAATGCCATTCGGCTCTGTTGCTTTATATTTGCTTTATGACACAGAG 2100
2041 ATGAGCTAAGCAATGCCATTCGGCTCTGTTGCTTTATATTTGCTTTATGACACAGAG 2100
2101 ATGTTTCTTCAAAAGTGGCAACACCAATATGCTTCTACTCTATGCTGGTGTGGTATCT 2160
2101 ATGTTTCTTCAAAAGTGGCAACACCAATATGCTTCTACTCTATGCTGGTGTGGTATCT 2160
2161 GTGTTGGTCTGTGGTTCGAGAGTATCCAGACCATGCGGGAAGAGATCTGACAC 2220
2161 GTGTTGGTCTGTGGTTCGAGAGTATCCAGACCATGCGGGAAGAGATCTGACAC 2220
2221 CGATCACACCTCTAGTGGCTTCAGTATGAATGGCATCTGCCCTCACTGTGGTATG 2280
2221 CGATCACACCTCTAGTGGCTTCAGTATGAATGGCATCTGCCCTCACTGTGGTATG 2280
2281 CATCAATATTTGGCTTCCCATCAGTACAAACATTTGTAAGTGGGCTCTGTTGTCTG 2340
2281 CATCAATATTTGGCTTCCCATCAGTACAAACATTTGTAAGTGGGCTCTGTTGTCTG 2340
2341 TTGGCTGGCTCCGGTCCAAGAGGCTGTGACTGGCGTCTCTTTCGTAACATTTTATGG 2400
2341 TTGGCTGGCTCCGGTCCAAGAGGCTGTGACTGGCGTCTCTTTCGTAACATTTTATGG 2400
2401 CCTGGTTGTACAGTCCCATTTCTGGAGTTATCAGTGGCTGCCATCATGGAATCTTCA 2460
2401 CCTGGTTGTACAGTCCCATTTCTGGAGTTATCAGTGGCTGCCATCATGGAATCTTCA 2460
2461 GATATGTCACTCTCAGAAATGTGAAGCTGTTGAGATTAATAATTTGTGCAATTTTGGGA 2520
2461 GATATGTCACTCTCAGAAATGTGAAGCTGTTGAGATTAATAATTTGTGCAATTTTGGGA 2520
2521 CCATCTTATAGGTATCTCTCCCTGAAAGATGATTACAGTGTAAACAGAGACTGACAA 2580
2521 CCATCTTATAGGTATCTCTCCCTGAAAGATGATTACAGTGTAAACAGAGACTGACAA 2580
2581 GAGTCTTTTATTTGGGAGCCAGAGGAGGAGTGTACTTGTGCTATAACTGCTTTTGT 2640
2581 GAGTCTTTTATTTGGGAGCCAGAGGAGGAGTGTACTTGTGCTATAACTGCTTTTGT 2640
2641 GCTAAATATGAATGCTCTCAAAATTTAGCTGTGTAAATAGCCCGGTTCCACTGGCTCT 2700
2641 GCTAAATATGAATGCTCTCAAAATTTAGCTGTGTAAATAGCCCGGTTCCACTGGCTCT 2700
2701 GCTGAGGTCCCTTCTTCTGGGCTGTGAATTCCTGTACATATTTCTACTTTTGTGA 2760
2701 GCTGAGGTCCCTTCTTCTGGGCTGTGAATTCCTGTACATATTTCTACTTTTGTGA 2760
2761 TCAGGCTCAATTTCCATTTATTTTAAATGTTGCTCTGAGAGTACTGTTGATTTTTTT 2820
2761 TCAGGCTCAATTTCCATTTATTTTAAATGTTGCTCTGAGAGTACTGTTGATTTTTTT 2820
2821 TCTTTTTTTTAAACCATGAAGAGCCGTTTGACAGAGCATGCTCTGCGTTGTTGGTTTCA 2880
2821 TCTTTTTTTTAAACCATGAAGAGCCGTTTGACAGAGCATGCTCTGCGTTGTTGGTTTCA 2880
2881 CAGCTTCTGCCCCCTCATGTCACAGGGAATTAACAACAAATAAATACTACAACTTCCCTT 2940
2881 CAGCTTCTGCCCCCTCATGTCACAGGGAATTAACAACAAATAAATACTACAACTTCCCTT 2940
2941 GTAGTCTCTTATATTAAGTAGAGTCTTGGTACTCTGCCCCCTCTGTCAGTAGTGGCAGGAT 3000
2941 GTAGTCTCTTATATTAAGTAGAGTCTTGGTACTCTGCCCCCTCTGTCAGTAGTGGCAGGAT 3000
3001 CTATTGGCATATTCGGAGCTTCTTAGAGGATGAGTCTTTTGAACACAGTGAATAATTT 3060
3001 CTATTGGCATATTCGGAGCTTCTTAGAGGATGAGTCTTTTGAACACAGTGAATAATTT 3060
3061 AAATTAGTAACTTTTTCGAAGCAGTTTATGACTGTTTATGCTTGAAGAGAGTGAAGAA 3120

Db 3061 AATTAGTAACCTTTTGGCAAGCAGTTATTGACTGTTATTGCTAAGAAAGTAAGAAA 3120
Qy 3121 GAAAAAGCCTGTTGGCAATCTTGGTATTATTTCTTTAAAGATTCTGGCAGTGTGGATGGAT 3180
Db 3121 GAAAAAGCCTGTTGGCAATCTTGGTATTATTTCTTTAAAGATTCTGGCAGTGTGGATGGAT 3180
Qy 3181 GAATGAAGTGAATGTGAATTTGGGCAAGTTAAATGGGACAGCCTTCAATGTTCAATTTG 3240
Db 3181 GAATGAAGTGAATGTGAATTTGGGCAAGTTAAATGGGACAGCCTTCAATGTTCAATTTG 3240
Qy 3241 TCTACCTCTTAACCTGAATTAAGAAAGCCTACAGTTTATAGAAAAAAGAAAA 3290
Db 3241 TCTACCTCTTAACCTGAATTAAGAAAGCCTACAGTTTATAGAAAAAAGAAAA 3290

RESULT 3
US-09-954-531-137
; Sequence 137, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 137
; LENGTH: 3220
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-137

Query Match 96.5%; Score 3175.2; DB 9; Length 3220;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3182; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 97 CCGACCGCGCGTGTGCGCGTGGCTCCAGCGCTGCGCGCTCGATCTCCTGCTCTCCG 156
Db 24 CCGGCGCGTGTGCGCGTGGCTCCAGCGCTGCGCGCTCGATCTCCTGCTCTCCG 83
Qy 157 CTCGCGCTCCCTTTTCCCTGGATGAATTTGGCTGCTTTCTCTCTCCGCCATGGAAATC 216
Db 84 CTCGCGCTCCCTTTTCCCTGGATGAATTTGGCTGCTTTCTCTCTCCGCCATGGAAATC 143
Qy 217 TGCTCCGCTGCTTTAGCCCTCTGAGCCAAAGAAACCCAGACAACAGATGCCATACGC 276
Db 144 TGCTCCGCTGCTTTAGCCCTCTGAGCCAAAGAAACCCAGACAACAGATGCCATACGC 203
Qy 277 AGCGTATAGCAGTAACCTCCAGCTCGGTTCTGTGCGGTAGTTACAGTATTAAATTTT 336
Db 204 AGCGTATAGCAGTAACCTCCAGCTCGGTTCTGTGCGGTAGTTACAGTATTAAATTTT 263
Qy 337 ATATAATATATATTTATATATAGCATTTTGTATACCTCATATCTGTTTACACATCTT 396
Db 264 ATATAATATATATTTATATATAGCATTTTGTATACCTCATATCTGTTTACACATCTT 323
Qy 397 GAAAGCGCTCAGTAGTTCTCTTACTAAACCAACCACTACTCCAGAGAATGGCAACCGTGA 456
Db 324 GAAAGCGCTCAGTAGTTCTCTTACTAAACCAACCACTACTCCAGAGAATGGCAACCGTGA 383
Qy 457 TTACCAAGTACTACAGCTGTACCGCGCTTCTGGTCTTGGTGGGACTACCTATGATGTC 516

Db 384 TTACCAAGTACTACAGCTGTACCGCGCTTCTGGTCTTTGGTGGACTACCTATGATGC 443
Qy 517 TCATCTCTGGCTTCATTATTGATTTGCTTCGCATTCTCCGTGGGAGCAATGATGTAG 576
Db 444 TCATCTCTGGCTTCATTATTGATTTGCTTCGCATTCTCCGTGGGAGCAATGATGTAG 503
Qy 577 CAAATTTCTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCCCTGAAGCAAGCCTGCATCC 636
Db 504 CAAATTTCTTTGGTACAGCTGTGGGCTCAGGTGTAGTGACCCCTGAAGCAAGCCTGCATCC 563
Qy 637 TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGCCAAAGTAGGCAAAACCA 696
Db 564 TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGCCAAAGTAGGCAAAACCA 623
Qy 697 TCCGGAAGGCTTTGATTGACGTGGAGATGTACAACCTCGACTCAAGGGCTACTGATGGCG 756
Db 624 TCCGGAAGGCTTTGATTGACGTGGAGATGTACAACCTCGACTCAAGGGCTACTGATGGCG 683
Qy 757 GCTCAGTCACTGCTATGTTTGGTCTGTGTGGCACTCGTGGCTTCGTTTTTGAAGC 816
Db 684 GCTCAGTCACTGCTATGTTTGGTCTGTGTGGCACTCGTGGCTTCGTTTTTGAAGC 743
Qy 817 TCCCTATTCTCGAAACCCATTGTTGTTGGTGGCACTATTGGTTCTCCCTCGTGGCAA 876
Db 744 TCCCTATTCTCGAAACCCATTGTTGTTGGTGGCACTATTGGTTCTCCCTCGTGGCAA 803
Qy 877 AGGGGAGGAGGCTGCAAGTGGTCTGAACCTGATGATAAAATTTGATGCTCTGTTCTGCT 936
Db 804 AGGGGAGGAGGCTGCAAGTGGTCTGAACCTGATGATAAAATTTGATGCTCTGTTCTGCT 863
Qy 937 CCCCACCTGCTTTCTGGAATTTATGTTGGAATTTATTTCTCTCTGTTCTGTCATTCATCC 996
Db 864 CCCCACCTGCTTTCTGGAATTTATGTTGGAATTTATTTCTCTCTGTTCTGTCATTCATCC 923
Qy 997 TCATAAGCAGATCCAGTTCTCTTAATGTTTGGAGCTTTGGCAGTTTCTATGCTCTGCA 1056
Db 924 TCATAAGCAGATCCAGTTCTCTTAATGTTTGGAGCTTTGGCAGTTTCTATGCTCTGCA 983
Qy 1057 CAGTTGGGAATAAACCTCTTTTCCATCATGTATCTGAGCACCCTGTTGCTGGGCTTTGACA 1116
Db 984 CAGTTGGGAATAAACCTCTTTTCCATCATGTATCTGAGCACCCTGTTGCTGGGCTTTGACA 1043
Qy 1117 AACTTCTCTGTGGGTACCATCTCATCTCGGTGGGATGTGCAGTTTCTGTGCCCTTA 1176
Db 1044 AACTTCTCTGTGGGTACCATCTCATCTCGGTGGGATGTGCAGTTTCTGTGCCCTTA 1103
Qy 1177 TCGTCTGTTTCTTTGTATGTCAGGATGAAGAGAAAAATTTGAAACGAGAAAAAAGTGTA 1236
Db 1104 TCGTCTGTTTCTTTGTATGTCAGGATGAAGAGAAAAATTTGAAACGAGAAAAAAGTGTA 1163
Qy 1237 GTCCCTTCTGAAAGCCCTTTAATGGAAAAAAGAAATAGCTTTGAAAGAAAGAACCATGAAGAAA 1296
Db 1164 GTCCCTTCTGAAAGCCCTTTAATGGAAAAAAGAAATAGCTTTGAAAGAAAGAACCATGAAGAAA 1223
Qy 1297 CAAAGTGTCTGTTGGTGAATTTGAAAAACAAGCATCTGTTTCTGAGGTAGGGCCTGCCA 1356
Db 1224 CAAAGTGTCTGTTGGTGAATTTGAAAAACAAGCATCTGTTTCTGAGGTAGGGCCTGCCA 1283
Qy 1357 CTGTGCCCTCCAGGCTGTGGTGGAGAGAGAAAGTCTCATTTCAAACTTGGAGATTTGG 1416
Db 1284 CTGTGCCCTCCAGGCTGTGGTGGAGAGAGAAAGTCTCATTTCAAACTTGGAGATTTGG 1343
Qy 1417 AGGAAGCTCCAGAGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGGAAACACCATAG 1476
Db 1344 AGGAAGCTCCAGAGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGGAAACACCATAG 1403
Qy 1477 ATAGCACCTGTAATGGTGCAGTGCAGTTGCTTAATGGAAACCTTTGTCCAGATTGAGTCAAG 1536
Db 1404 ATAGCACCTGTAATGGTGCAGTGCAGTTGCTTAATGGAAACCTTTGTCCAGATTGAGTCAAG 1463
Qy 1537 CCGTCAAGCAACCAATAAATCACTCAGTGGCCACTACAGTATCACACCTGTCATAGGATT 1596

1464 CCCTCAGCAACCAATAAATCCAGTGGCCACTCCAGTATCACACCGTGCAATGAGATT 1523
QY
1597 CCGGCTGTACAAGAGCTACTCCATAAATTACATCTTGCCAGGTTGGGAGATTGCATGG 1656
Db
1524 CCGGCTGTACAAGAGCTACTCCATAAATTACATCTTGCCAAAGTTGGGAGATTGCATGG 1583
QY
1657 GAGACTCCGGTGACAAACCCCTTAAGGCGCAATAATAGCTATACCTTATACCATGGCAA 1716
Db
1584 GAGACTCCGGTGACAAACCCCTTAAGGCGCAATAATAGCTATACCTTATACCATGGCAA 1643
QY
1717 TATGTGGCATGCTCTGGATTCATCCGTGCCAAAGAGTGAAACAGAGGGCGAAGAAA 1776
Db
1644 TATGTGGCATGCTCTGGATTCATCCGTGCCAAAGAGTGAAACAGAGGGCGAAGAAA 1703
QY
1777 TGAAGAAGCTGACATGGCCTAATGACAGCTCCAAAGAGCGAATTCGAATGGACAGTTACA 1836
Db
1704 TGAAGAAGCTGACATGGCCTAATGACAGCTCCAAAGAGCGAATTCGAATGGACAGTTACA 1763
QY
1837 CCAGTTACTGCAATGCTGTCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA 1896
Db
1764 CCAGTTACTGCAATGCTGTCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA 1823
QY
1897 AGCAGAGATGGGTCTAGGTGACAGAAAGAAAGTAATGGCTCTCTAGAAAGATGGTATG 1956
Db
1824 AGCAGAGATGGGTCTAGGTGACAGAAAGAAAGTAATGGCTCTCTAGAAAGATGGTATG 1883
QY
1957 ACCAGNATAGCCTGAAGTCTCTCTCTCTTCCAGTTCCTGACAGATCCTTACAGCCTGCT 2016
Db
1884 ACCAGNATAGCCTGAAGTCTCTCTCTCTTCCAGTTCCTGACAGATCCTTACAGCCTGCT 1943
QY
2017 TTGGGTCAATTCGCCCATGGTGGCAATGACATGAAGCAATGCAATGGGCCCTCTGGTTGCTT 2076
Db
1944 TTGGGTCAATTCGCCCATGGTGGCAATGACATGAAGCAATGCAATGGGCCCTCTGGTTGCTT 2003
QY
2077 TATATTGGTTTATGACACAGAGATGTTTCTTCAAAAGTGGAACACCAATATGGCTTC 2136
Db
2004 TATATTGGTTTATGACACAGAGATGTTTCTTCAAAAGTGGAACACCAATATGGCTTC 2063
QY
2137 TACTATGTTGGTGGTATCTGTGTTGGTCTGTGGTGGTGGGAGAGAGTTATCC 2196
Db
2064 TACTATGTTGGTGGTATCTGTGTTGGTCTGTGGTGGTGGGAGAGAGTTATCC 2123
QY
2197 AGACCATGGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTCAATTAAGAACTGG 2256
Db
2124 AGACCATGGGGAAGGATCTGACACCGATCACACCTCTAGTGGCTTCAATTAAGAACTGG 2183
QY
2257 CATCTGCCCTCAGTGGTGAATTCATCAAAATATGGCCTTCCCATCAGTACAAACATTT 2316
Db
2184 CATCTGCCCTCAGTGGTGAATTCATCAAAATATGGCCTTCCCATCAGTACAAACATTT 2243
QY
2317 GTAAAGTGGGCTCTGTGTTGTTGGTGGCTCCGGTCCAGAGGCTGTTGACTGGC 2376
Db
2244 GTAAAGTGGGCTCTGTGTTGTTGGTGGCTCCGGTCCAGAGGCTGTTGACTGGC 2303
QY
2377 GTCTCTTCCGTAACATTTTATGGCCTGGTGTGTGACAGTCCCATTTCTGGAGTTATCA 2436
Db
2304 GTCTCTTCCGTAACATTTTATGGCCTGGTGTGTGACAGTCCCATTTCTGGAGTTATCA 2363
QY
2437 GTGCTGCCATCAGGCAATCTTCAAGATATGTCATCTCAGAAATGTGAAGCTGTTGAGAT 2496
Db
2364 GTGCTGCCATCAGGCAATCTTCAAGATATGTCATCTCAGAAATGTGAAGCTGTTGAGAT 2423
QY
2497 TAAATTTGTGCAATGTTTGGGACCATCTAGGTATTCCTGCTCCCTGCAAGATGATT 2556
Db
2424 TAAATTTGTGCAATGTTTGGGACCATCTAGGTATTCCTGCTCCCTGCAAGATGATT 2483
QY
2557 ACAGTGTAAACAGAAAGCTGACAAAGCTCTTTTATTTGGGAGCCAGAGAGGGAAGTGT 2616
Db
2484 ACAGTGTAAACAGAAAGCTGACAAAGCTCTTTTATTTGGGAG-CAGAGAGGGAAGTGT 2542
QY
2617 TACTTGTGCTATAACCTGTTTGTGTAATAATGAATTTGCTCAAAATTAGCTGTGTAAA 2676
Db
2543 TACTTGTGCTATAACCTGTTTGTGCTAAATATGAATTTGCTCAAAATTAGCTGTGTAAA 2602

2677 ATAGCCCGGTTCCACTGGCTCCTGCTGAGGTGCCCTTTCTTTCTGGGCTGTGAATTCCT 2736
Db
2603 ATAGCCCGGTTCCACTGGCTCCTGCTGAGGTGCCCTTTCTTTCTGGGCTGTGAATTCCT 2662
QY
2737 GTACATATTTCTCTACTTTTGTATCAGGCTTCAATCCATATATGTTTAAATGTTGCTC 2796
Db
2663 GTACATATTTCTCTACTTTTGTATCAGGCTTCAATCCATATATGTTTAAATGTTGCTC 2722
QY
2797 TGAAGATGACTGTGATTTTCTTTTCTTTTAAACCATGAAGCGCTTTGACAGAG 2856
Db
2723 TGAAGATGACTGTGATTTTCTTTTCTTTTAAACCATGAAGCGCTTTGACAGAG 2782
QY
2857 CATGCTCTGCGTGTGTTTCCACAGCTTCTGCCCTCACATGACACAGGATTTAAACAAC 2916
Db
2783 CATGCTCTGCGTGTGTTTCCACAGCTTCTGCCCTCACATGACACAGGATTTAAACAAC 2842
QY
2917 AAAAAATAAACTACAACTTCCCTGTAGTCTCTTATATAAGTAGAGTCTTGGTACTCTG 2976
Db
2843 AAAAAATAAACTACAACTTCCCTGTAGTCTCTTATATAAGTAGAGTCTTGGTACTCTG 2902
QY
2977 CCCTCTGTCTAGTAGTGGCAGATCTATTTGGCATATTCGGAGCTTCTTAGAGGATGAG 3036
Db
2903 CCCTCTGTCTAGTAGTGGCAGATCTATTTGGCATATTCGGAGCTTCTTAGAGGATGAG 2962
QY
3037 GTTCTTTGAAACACAGTGAATAATTTAAATTTAGTAACTTTTTCGAAGCAGTTTATTGACTG 3096
Db
2963 GTTCTTTGAAACACAGTGAATAATTTAAATTTAGTAACTTTTTCGAAGCAGTTTATTGACTG 3022
QY
3097 TTATTGCTAAGAAAGTAAGAAAGAAAGAAAGCTTGTGGCAATCTTGGTATTCTTTTAA 3156
Db
3023 TTATTGCTAAGAAAGTAAGAAAGAAAGAAAGCTTGTGGCAATCTTGGTATTCTTTTAA 3082
QY
3157 GATTCTTGGCAGTGTGGATGGAATGAAGTGAATGTGAACCTTTGGGCAAGTTAAAT 3216
Db
3083 GATTCTTGGCAGTGTGGATGGAATGAAGTGAATGTGAACCTTTGGGCAAGTTAAAT 3142
QY
3217 GGGCAGAGCTTCCATGTTCTTCACTTCTTAACTGAATAAAAAGCCCTACAGTTTTT 3276
Db
3143 GGGCAGAGCTTCCATGTTCTTCACTTCTTAACTGAATAAAAAGCCCTACAGTTTTT 3202
QY
3277 TAGAAAAAA 3285
Db
3203 TAGAAAAAA 3211

RESULT 4

US-09-954-531-356
; Sequence 356, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can-
; FILE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 356
; LENGTH: 3220
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-954-531-356

Query Match 96.5%; Score 3175.2; DB 9; Length 3220;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3182; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY	97	CCGACCGGCGGTGTCGGTGGCTCCAGCCGCTGCGGCTCGATCTCTGTCCTCCG 156
DB	24	CGGGCGGTGCGGTGTCGGTGGCTCCAGCCGCTGCGGCTCGATCTCTGTCCTCCG 83
QY	157	CTCCGCCCTCCCTTTTCCCTGGATGAACCTTGCGTCTCTTCTCTCCGCCATGGAATTC 216
DB	84	CTCCGCCCTCCCTTTTCCCTGGATGAACCTTGCGTCTCTTCTCTCCGCCATGGAATTC 143
QY	217	TGCTCCGCTGCTTTAGCCCTCTTGAGCCAAAGAAACCCAGACAAAGATGCCATACGC 276
DB	144	TGCTCCGCTGCTTTAGCCCTCTTGAGCCAAAGAAACCCAGACAAAGATGCCATACGC 203
QY	277	AGCGTATAGCAGTAACCTCCCGAGCTGGTTCTGTCGCGTAGTTTACAGTATTAATTTT 336
DB	204	AGCGTATAGCAGTAACCTCCCGAGCTGGTTCTGTCGCGTAGTTTACAGTATTAATTTT 263
QY	337	ATATAATATATATTTATTTATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTT 396
DB	264	ATATAATATATATTTATTTATAGCATTTTGTATACCTCATATTTCTGTTTACACATCTT 323
QY	397	GAAAGCGCTCAGTAGTTCTCTTACTAAACAAACCACTACTCCAGAGAAATGGCAACGGTGA 456
DB	324	GAAAGCGCTCAGTAGTTCTCTTACTAAACAAACCACTACTCCAGAGAAATGGCAACGGTGA 383
QY	457	TTACCAAGTACTACAGCTGCTACCGCGCTTCTGGTCTTGTGGTGGACTACTATGATGTC 516
DB	384	TTACCAAGTACTACAGCTGCTACCGCGCTTCTGGTCTTGTGGTGGACTACTATGATGTC 443
QY	517	TCATCTCGGCTTCTATTATTTGCAATTTGCTTGGCAATTTCTCCGGGGGCAATGATGTAG 576
DB	444	TCATCTCGGCTTCTATTATTTGCAATTTGCTTGGCAATTTCTCCGGGGGCAATGATGTAG 503
QY	577	CAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGAACCTTGAAGCAAGCCTGCATCC 636
DB	504	CAAAATCTTTTGGTACAGCTGTGGGCTCAGGTGTAGTGAACCTTGAAGCAAGCCTGCATCC 563
QY	637	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCAAAAGTAGGCAACCA 696
DB	564	TAGCTAGCATCTTTGAAACAGTGGGCTCTGTCTTACTGGGGGCAAAAGTAGGCAACCA 623
QY	697	TCCGGAAGGCTTGATTGAGCTGGAGTGTACAACTCGACTCAAGGCTACTGATGGCCG 756
DB	624	TCCGGAAGGCTTGATTGAGCTGGAGTGTACAACTCGACTCAAGGCTACTGATGGCCG 683
QY	757	GCTCAGTCAAGTCTATGTTGGTCTGCTGTGGCAACTCGTGGCTTCGTTTTCGAAGC 816
DB	684	GCTCAGTCAAGTCTATGTTGGTCTGCTGTGGCAACTCGTGGCTTCGTTTTCGAAGC 743
QY	817	TCCTATTTCTGGAACCAATTTGATTTGTTGGTGCAACTATTTGTTTCTCCCTCGTGCA 876
DB	744	TCCTATTTCTGGAACCAATTTGATTTGTTGGTGCAACTATTTGTTTCTCCCTCGTGCA 803
QY	877	AGGGGAGGAGGGGTCAAGTGGTCTGAACTGATTAATAATTTGATGCTTGGTTCGTGT 936
DB	804	AGGGGAGGAGGGGTCAAGTGGTCTGAACTGATTAATAATTTGATGCTTGGTTCGTGT 863
QY	937	CCCCACTGCTTTCTGGAATTTATGCTGGAATTTTATTTCTTCTGGTTCGTGCAATTCATCC 996
DB	864	CCCCACTGCTTTCTGGAATTTATGCTGGAATTTTATTTCTTCTGGTTCGTGCAATTCATCC 923
QY	997	TCCATTAAGCAGATCCAGTTCTTAATGGTTTGGAGCTTTGCCAGTTTCTATGCTGCA 1056
DB	924	TCCATTAAGCAGATCCAGTTCTTAATGGTTTGGAGCTTTGCCAGTTTCTATGCTGCA 983
QY	1057	CAGTTGGAATTAACCTCTTTTCCATCATGTATACTGGAGCACCGTTCCTGGGCTTTGACA 1116
DB	984	CAGTTGGAATTAACCTCTTTTCCATCATGTATACTGGAGCACCGTTCCTGGGCTTTGACA 1043

QY	1117	AACCTTCCTCTGTGGGTACCATCTCATCTCGTGGGATGTGCAGTCTTCTGTGCCCTTA 1176
DB	1044	AACCTTCCTCTGTGGGTACCATCTCATCTCGTGGGATGTGCAGTCTTCTGTGCCCTTA 1103
QY	1177	TGCTCTGGTCTTGTATGTCCAGGATGAAGAGAAAAATTTGAACGAGAAAAATAAGTGTA 1236
DB	1104	TGCTCTGGTCTTGTATGTCCAGGATGAAGAGAAAAATTTGAACGAGAAAAATAAGTGTA 1163
QY	1237	GTCTCTTCTGAAAGCCCTTTAATGGAAAAAAGATAGCTTTGAAAGAGAACCATGAAGAAA 1296
DB	1164	GTCTCTTCTGAAAGCCCTTTAATGGAAAAAAGATAGCTTTGAAAGAGAACCATGAAGAAA 1223
QY	1297	CAAAAGTGTCTGTGGTGTATTTGAAAAAACAAGCATCTCTGTTTCTTGAGGTAGGSCCTGCCA 1356
DB	1224	CAAAAGTGTCTGTGGTGTATTTGAAAAAACAAGCATCTCTGTTTCTTGAGGTAGGSCCTGCCA 1283
QY	1357	CTGTGCCCTCCAGGCTGTGGTGGAGAGAACAGTCTCATTTCAAACCTTGGAGATTTGG 1416
DB	1284	CTGTGCCCTCCAGGCTGTGGTGGAGAGAACAGTCTCATTTCAAACCTTGGAGATTTGG 1343
QY	1417	AGGAAGCTCCAGAGAGAGAGGCTTCCAGCGTGGACTTGAAGAGGAAAAACAGCATAG 1476
DB	1344	AGGAAGCTCCAGAGAGAGAGGCTTCCAGCGTGGACTTGAAGAGGAAAAACAGCATAG 1403
QY	1477	ATAGCACCGTGAATGGTGCAGTGCAGTTGCCCTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1536
DB	1404	ATAGCACCGTGAATGGTGCAGTGCAGTTGCCCTAATGGGAACCTTGTCCAGTTTCAGTCAAG 1463
QY	1537	CGGTGAGCAACCAATAAATCACTCCAGTGGCACTACACAGTATCACACCGTGCAATAGGATT 1596
DB	1464	CGGTGAGCAACCAATAAATCACTCCAGTGGCACTACACAGTATCACACCGTGCAATAGGATT 1523
QY	1597	CCGGCTGTACAAAGAGCTACTCCATAAATTAACCTTCCCAAGGTGGGAGATTGCATGG 1656
DB	1524	CCGGCTGTACAAAGAGCTACTCCATAAATTAACCTTCCCAAGGTGGGAGATTGCATGG 1583
QY	1657	GAGACTCCGGTGACAAACCTTTAAGCGCAATAAATAGCTATATCTTCTATACCATGGCAA 1716
DB	1584	GAGACTCCGGTGACAAACCTTTAAGCGCAATAAATAGCTATATCTTCTATACCATGGCAA 1643
QY	1717	TATGTGGCATGCTCTGGAATTCATTTCCGTGCCAAAGAGGTGAAACGAAGGGGGAAGAAA 1776
DB	1644	TATGTGGCATGCTCTGGAATTCATTTCCGTGCCAAAGAGGTGAAACGAAGGGGGAAGAAA 1703
QY	1777	TGGAAGAGCTGACATGGCTTAATGAGACTCCAAAGAGCGAATTCGAATGGACAGTTACA 1836
DB	1704	TGGAAGAGCTGACATGGCTTAATGAGACTCCAAAGAGCGAATTCGAATGGACAGTTACA 1763
QY	1837	CCAGTTTACTGCAATGCTGTCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA 1896
DB	1764	CCAGTTTACTGCAATGCTGTCTGACCTTCACTCAGCATCTGAGATAGACATGAGTGTCA 1823
QY	1897	AGGAGAGATGGGTCTAGGTGACAGAAAGAAAGTAATGGCTCTCTAGAAGAAATGGTATG 1956
DB	1824	AGGAGAGATGGGTCTAGGTGACAGAAAGAAAGTAATGGCTCTCTAGAAGAAATGGTATG 1883
QY	1957	ACGAGATAAGCTCAAGTCTCTCTCTCCAGTTCCTGAGATCCCTTACAGCTGCT 2016
DB	1884	ACGAGATAAGCTCAAGTCTCTCTCTCCAGTTCCTGAGATCCCTTACAGCTGCT 1943
QY	2017	TTGGGTTCATTTCGCCCATGGTGGCAATGACGTAAAGCAATGCCATTGGGCTCTGGTTGCTT 2076
DB	1944	TTGGGTTCATTTCGCCCATGGTGGCAATGACGTAAAGCAATGCCATTGGGCTCTGGTTGCTT 2003
QY	2077	TATATTTGGTTTATGACACAGAGATGTTTCTTCAAAAAGTGGCAACACCAATATGGCTTC 2136
DB	2004	TATATTTGGTTTATGACACAGAGATGTTTCTTCAAAAAGTGGCAACACCAATATGGCTTC 2063
QY	2137	TACTCTATGGTGGTCTGTTGTTCTGTTGGTCTGTTGGTCTGTTGGGGAAGAGTTATCC 2196
DB	2064	TACTCTATGGTGGTCTGTTGTTCTGTTGGTCTGTTGGTCTGTTGGGGAAGAGTTATCC 2123

Qy	2197	AGACCATGGGGAAGGATCTGACACCGATCAACCCCTCTAGTGGCTTCAGTATTGAAC	2255
Db	2124	AGACCATGGGGAAGGATCTGACACCGATCAACCCCTCTAGTGGCTTCAGTATTGAAC	2183
Qy	2257	CATCTGCCCTCAGTCTGGTGAATGTCATCAAAATATGGCCCTCCCATCAGTACAAACAT	2316
Db	2184	CATCTGCCCTCAGTCTGGTGAATGTCATCAAAATATGGCCCTCCCATCAGTACAAACAT	2243
Qy	2317	GTAAGTGGGCTCTGTTGTGCTCTGGCTGGCTCCGGTCCAAAGAGGCTGTGTGACTGGC	2376
Db	2244	GTAAGTGGGCTCTGTTGTGCTCTGGCTGGCTCCGGTCCAAAGAGGCTGTGTGACTGGC	2303
Qy	2377	GTCTCTTTGCPAACATTTTTATAGCCCTGGTTGTGACAGTCCCATTTCTGGAGTTATCA	2436
Db	2304	GTCTCTTTGCPAACATTTTTATAGCCCTGGTTGTGACAGTCCCATTTCTGGAGTTATCA	2363
Qy	2437	GTGCTGCCATCATGGCAATCTTCAGATATGTCATCTCCAGAAATGTAAGCTGTTGAGAT	2496
Db	2364	GTGCTGCCATCATGGCAATCTTCAGATATGTCATCTCCAGAAATGTAAGCTGTTGAGAT	2423
Qy	2497	TAAATTTGTGTCAAATGTTGGGACCACTTTAGGTATTTCTGCTCCCTCAAGAATGATT	2556
Db	2424	TAAATTTGTGTCAAATGTTGGGACCACTTTAGGTATTTCTGCTCCCTCAAGAATGATT	2483
Qy	2557	ACAGTGTAAACAGAAGACTGACAAGAGTCTTTTTTATTTGGGAGCCAGAGAGGGAAGTGT	2616
Db	2484	ACAGTGTAAACAGAAGACTGACAAGAGTCTTTTTTATTTGGGAG - CAGAGAGGGAAGTGT	2542
Qy	2617	TACTTGTGCTATAACTGCTTTTGTGCTAAATATGAAATGTCTCAAATTAGCTGTGTAAA	2676
Db	2543	TACTTGTGCTATAACTGCTTTTGTGCTAAATATGAAATGTCTCAAATTAGCTGTGTAAA	2602
Qy	2677	ATAGCCGGGTTCCACTGGCTCCGCTGAGGTCCCTTCTCTCTGGGCTGTGAATTCCT	2736
Db	2603	ATAGCCGGGTTCCACTGGCTCCGCTGAGGTCCCTTCTCTCTGGGCTGTGAATTCCT	2662
Qy	2737	GTACATATTCTCTACTTTTTTGTATCAGGCTTCAAATTCATATGTTTTTAATGTGTCTC	2796
Db	2663	GTACATATTCTCTACTTTTTTGTATCAGGCTTCAAATTCATATGTTTTTAATGTGTCTC	2722
Qy	2797	TGAAGATGACTGTGATTTTTTTTTTTTTTTTTTAAACCATGAAGAGCCGTTTGACAGAG	2856
Db	2723	TGAAGATGACTGTGATTTTTTTTTTTTTTTTTTAAACCATGAAGAGCCGTTTGACAGAG	2782
Qy	2857	CATGCTCTGCGTTGTGGTTTCCACAGCTCTGCGCTCCATGACAGAGGATTTAAACAAC	2916
Db	2783	CATGCTCTGCGTTGTGGTTTCCACAGCTCTGCGCTCCATGACAGAGGATTTAAACAAC	2842
Qy	2917	AAAAATATACTACAACCTCCCTGTAGTCTCTTATATAGTAGAGTCTTGGTACTCTG	2976
Db	2843	AAAAATATACTACAACCTCCCTGTAGTCTCTTATATAGTAGAGTCTTGGTACTCTG	2902
Qy	2977	CCCTCTGTGAGTAGTGGCAGGATCTATTTGGCATATTTGGGAGGCTCTTAGAGGGATGAG	3036
Db	2903	CCCTCTGTGAGTAGTGGCAGGATCTATTTGGCATATTTGGGAGGCTCTTAGAGGGATGAG	2962
Qy	3037	GTTCTTTGAACACAGTGAATAATTTAAATAGTAACTTTTTTGAAGCAGATTTATGACTG	3096
Db	2963	GTTCTTTGAACACAGTGAATAATTTAAATAGTAACTTTTTTGAAGCAGATTTATGACTG	3022
Qy	3097	TTATTGCTAAGAAGAAGTAAGAAAGAAAAAGCCCTGTTGGCAATCTTGGTTATTTCTTTAA	3156
Db	3023	TTATTGCTAAGAAGAAGTAAGAAAGAAAAAGCCCTGTTGGCAATCTTGGTTATTTCTTTAA	3082
Qy	3157	GATTTCTGGCAGTGTGGGATGGATGAATGAAGTGGAAATGTGAACTTTGGGCAAGTTAAAT	3216
Db	3083	GATTTCTGGCAGTGTGGGATGGATGAATGAAGTGGAAATGTGAACTTTGGGCAAGTTAAAT	3142
Qy	3217	GGGACAGCCTTCCATGTTTCATTTCTCTACCTCTTAAGTGAATAAAAAGCCTACAGTTTT	3276
Db	3143	GGGACAGCCTTCCATGTTTCATTTCTCTACCTCTTAAGTGAATAAAAAGCCTACAGTTTT	3202
Qy	3277	TAGAAAAAA 3285	

Db 3203 TAGAAAAA 3211

RESULT 5

```

US-09-879-536-623
; Sequence 623, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENE
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CDD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/089,
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version
; SEQ ID NO 623
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(662)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-536-623

```

Query Match	15.1%	Score 495.2	DB 10	Length 662
Best Local Similarity	90.2%	Pred. No. 6.7e-134		
Matches 591	Conservative 0	Mismatches 54	Indels 10	Gaps 6
Qy	1606	ACAAGAGCTACTCCATAAATTACATCTTGCCAAAGTGGGAGATTGCATGGGAGACTCCG	1665	
Db	1	ACAAAGAGCTACTCCATAAATTACATCTTGCCAAAGTGGGAGATTGCATGGGAGACTCCG	60	
Qy	1666	GTGACAAACCTTAAAGCGGCAATAATAGCTATACTTCTTATACCATGGCNAATATGTGGCA	1725	
Db	61	GTGACAAACCTTAAAGCGGCAATAATAGCTATACTTCTTATACCATGGCNAATATGTGGCA	120	
Qy	1726	TGCTCTCGGAATTCATTCCGTGCCAAAGAGGTCGAAACAGAAAGGCGCAAGAAATCGAGAACG	1785	
Db	121	TGCTCTCGGAATTCATTCCGTGCCAAAGAGGTCGAAACAGAAAGGCGCAAGAAATCGAGAACG	180	
Qy	1786	TGACATGGCCTAATGACAGACTCCAAGAAAGCGAAATTCGAATGGACAGTTACACCAAGTTACT	1845	
Db	181	TGACATGGCCTAATGCGGACTCCAAGAAAGCGAAATTCGAATGGACAGTTACACCAAGTTACT	240	
Qy	1846	GCAATGCTGTGCTGACCTTTCACTCAGCATCTCAGATAGACATGAGTGTCAAGGCGAGAGA	1905	
Db	241	GCAATGCTGTGCTGACCTTTCACTCAGCATCTCAGATAGACATGAGTGTCAAGGCGAGAGA	300	
Qy	1906	TGGGCTTAGGTGACAGAAAAAGGAA--GTAAATGGCTCTCTAGAAAGAAATGGTATGACCAGGA	1963	
Db	301	TGGGCTTAGGTGACAGAAAAAGGAAATTAATGGGCTCTCTAGAAAGAAATGGTATGACCAGGA	360	
Qy	1964	TAGCCCTGAGTCTCTCTCTCTCTCCAGTTCCCTGCGAGATCCCTTACAGCCCTGCTTTGGGTC	2023	
Db	361	TAGCCCTGAAAGTCTCTCTCTCTCTCTTCAGTTCTCTGCGAGANCCCTTACAGCCCTGCTTTGGGTC	420	
Qy	2024	ATTTCGCCCATGGTGGCAATGACGTAAAGCAATAGCCA--TTGGGCGCTCTGGTTGCTTTTATATT	2082	

Db 421 ATTGGCCCATGGTGGCAATGACGTAAAGCCATGCAATTTGGGCTCTGGGTGCTTATATT 480
QY 2083 TGGTTTATGACACAGGAGATGTTTCTTCAAAAGTGGCAACACCAATA-TGGCTTCTACTC 2141
Db 481 TGGGTTATGACCCNNGAGANG-TTCTTCAAAAGTGGCAACACCAATATTGNTTCTACTC 539
QY 2142 TAT--GGTGTGTTGTTATCTGTGTTGG----TCTGTGGGTTTGGGGAAGAAGTATTATCC 2196
Db 540 TANTGGNGGGGTGGGATCTGNGGTTGTCTGTGGGGTTTGGGGAAGAAAAAGTTTTC 599
QY 2197 AGACCATGGGAAGGATCTGACACCGCATCACCCCTCTAGTGGCTTCAGTATTGA 2251
Db 600 CNACCTTGGGAAGGATTGGCCNCCGTACACCCCTTTAAGGGTTTNGTATTTGA 654

RESULT 6

US-09-867-701-2803
; Sequence 2803, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2803
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-701-2803

Query Match 14.0%; Score 459.8; DB 10; Length 463;
Best Local Similarity 99.6%; Pred. No. 1.1e-123;
Matches 461; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2828 TTATACCATGAGAGCCGTTTACACAGAGCATCTCTGGTGTGTTGGTTTCCACGCTTC 2887
Db 1 TTTAAACCATGAAGAGCCGTTTACACAGAGCATCTCTGGTGTGTTGGTTTCCACGCTTC 60
QY 2888 TGCCTCTCACATGCACAGGAGTTTAAACAAACAAATATATACTACAACTTCCCTTGTAGTCT 2947
Db 61 TGCCTCTCACATGCACAGGAGTTTAAACAAACAAATATATACTACAACTTCCCTTGTAGTCT 120
QY 2948 CTTATATAAGTAGAGTCTCTTGGTACTCTGCCCTCTCTGTCTAGTAGTGGCAGGATCTATTGG 3007
Db 121 CTTATATAAGTAGAGTCTCTTGGTACTCTGCCCTCTCTGTCTAGTAGTGGCAGGATCTATTGT 180
QY 3008 CATATTGGGAGCTCTTAGAGGGATGAGTCTCTTTGAACACAGTGAATAATTAAATTAG 3067
Db 181 CATATTGGGAGCTCTTAGAGGGATGAGTCTCTTTGAACACAGTGAATAATTAAATTAG 240
QY 3068 TAACTTTTGTGCAAGCAGTTTATTGACTGTATTGCTTGAACAGAGTGAAGAAAGAAAAG 3127
Db 241 TAACTTTTGTGCAAGCAGTTTATTGACTGTATTGCTTGAACAGAGTGAAGAAAGAAAAG 300
QY 3128 CTTGTTGGCAATCTTGGTTATTCTTTTAAAGATTCTGGCAGTGTGGGATGGATGAATGAA 3187
Db 301 CTTGTTGGCAATCTTGGTTATTCTTTTAAAGATTCTGGCAGTGTGGGATGGATGAATGAA 360
QY 3188 GTGGAATGGAATCTTGGGCAAGTTAAATGGGACAGCTTCCATGTTCATTGTTCTACCT 3247
Db 361 GTGGAATGGAATCTTGGGCAAGTTAAATGGGACAGCTTCCATGTTCATTGTTCTACCT 420
QY 3248 CTTAACTGAATAAAAGCCTACAGTTTATAGAAAAA 3290
Db 421 CTTAACTGAATAAAAGCCTACAGTTTATAGAAAAA 463

RESULT 7

US-09-879-536-521
; Sequence 521, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Buehnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 521
; LENGTH: 613
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)-(613)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-536-521

Query Match 13.7%; Score 451.4; DB 10; Length 613;
Best Local Similarity 99.2%; Pred. No. 4.1e-121;
Matches 474; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1131 GGTACATCTCTCATCTCGGTGGGATGTGCAAGTTTCTGTGCCCTTATCGTCTGGTCTTT 1190
Db 1 GGTACATCTCTCATCTCGGTGGGATGTGCAAGTTTCTGTGCCCTTATCGTCTGGTCTTT 60
QY 1191 GTATGTCCAGATGAAGAGAAAATTTGAACAGAAAATAAGTGTAGTCTCTGAAAGC 1250
Db 61 GTATGTCCAGATGAAGAGAAAATTTGAACAGAAAATAAGTGTAGTCTCTGAAAGC 120
QY 1251 CCCTTAATGAAAAAAGAAATAGCTTTGAAAGAGACCATGAAGAAAACAAAGTTGTCTGTT 1310
Db 121 CCCTTAATGAAAAAAGAAATAGCTTTGAAAGAGACCATGAAGAAAACAAAGTTGTCTGTT 180
QY 1311 GGTGATATTGAAAAACAGCATCTCTGTTCTGTAGGTAGGCGCTGCCACTGTGCCCTCCAG 1370
Db 181 GGTGATATTGAAAAACAGCATCTCTGTTCTGTAGGTAGGCGCTGCCACTGTGCCCTCCAG 240
QY 1371 GCTGTGGTGGAGAGAGACAGTCTCATTTCAAACTTGGAGATTTGAGAGAGCTCCAGAG 1430
Db 241 GCTGTGGTGGAGAGAGAACAGTCTCATTTCAAACTTGGAGATTTGAGAGAGCTCCAGAG 300
QY 1431 AGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGGAAAACAGCATAGATAGCACCGTGAAT 1490
Db 301 AGAGAGAGGCTTCCAGCGTGGACTTTGAAAGAGGAAAACAGCATAGATAGCACCGTGAAT 360
QY 1491 GGTGAGTGCAGTGTGCTTAATGGGAACCTTGTCCAGTTCAGTTC-AAGCGTCAGCAACCA 1549
Db 361 GGTGAGTGCAGTGTGCTTAATGGGAACCTTGTCCAGTTCAGTTCAGTTCAGCAACCA 420
QY 1550 AATAAATCCAGTGGCCACTACAGTATCACACCGTGCAT-TAGGATTCGGGCTGTA 1606
Db 421 AATAAATCCAGTGGCCACTACAGTATCACACCGTGCATTAAGGATTCGGGCTGTA 478

RESULT 8

US-09-920-300A-1304/c
; Sequence 1304, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1304
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 454
; OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-1304

Query Match 13.6%; Score 448; DB 10; Length 495;
Best Local Similarity 99.4%; Pred. No. 3.4e-120;
Matches 460; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1102 TGCTGGGCTTTGACAACTTCCTCTGTGGGGTACCATCCTCATCTCGGTGGGATGTCAG 1161
DB 461 TGCTGGGNTTTGACAACTTCCTCTGTGGGGTACCATCCTCATCTCGGTGGGATGTCAG 402
QY 1162 TTTTCTGTGCCCTTATCGTCTGGTCTTTGTATGTCCAGGATGAAGAGAAAAATTGAAC 1221
DB 401 TTTTCTGTGCCCTTATCGTCTGGTCTTTGTATGTCCAGGATGAAGAGAAAAATTGAAC 342
QY 1222 GAGAAATAAGTGTAGTCTTCTGAAAGCCCTTAAATGAAAAAAGAAATAGCTTGAAG 1281
DB 341 GAGAAATAAGTGTAGTCTTCTGAAAGCCCTTAAATGAAAAAAGAAATAGCTTGAAG 282
QY 1282 AAGACCATGAAGAAACAAAGTTGTCTGTGGTGATATTGAAACAAAGCATCTGTTCTG 1341
DB 281 AAGACCATGAAGAAACAAAGTTGTCTGTGGTGATATTGAAACAAAGCATCTGTTCTG 222
QY 1342 AGGTAGGGCTGCCACTGTGCCCTCCAGGCTGTGGTGAGGAGAGAACAGTCTCATTTCA 1401
DB 221 AGGTAGGGCTGCCACTGTGCCCTCCAGGCTGTGGTGAGGAGAGAACAGTCTCATTTCA 162
QY 1402 AACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGGCTTCCAGGCTGGACTTGAAG 1461
DB 161 AACTTGGAGATTTGGAGGAAGCTCC--AGAGAGAGAGGCTTCCAGGCTGGACTTGAAG 104
QY 1462 AGGAAACACATAGATAGCAGCGTGAATGGTGCAGTTCAGTTCCTTAATGGAACTTG 1521
DB 103 AGGAAACACATAGATAGCAGCGTGAATGGTGCAGTTCAGTTCCTTAATGGAACTTG 44
QY 1522 TCCAGTTTCACTCAAGCCGTGAGCAACCAAAATAAATCCAGTGG 1564
DB 43 TCCAGTTTCACTCAAGCCGTGAGCAACCAAAATAAATCCAGTGG 1

RESULT 9

US-10-033-528-1304/c
; Sequence 1304, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1304
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 454
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-1304
Query Match 13.6%; Score 448; DB 12; Length 495;
Best Local Similarity 99.4%; Pred. No. 3.4e-120;
Matches 460; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1102 TGCTGGGCTTTGACAACTTCCTCTGTGGGGTACCATCCTCATCTCGGTGGGATGTCAG 1161
DB 461 TGCTGGGNTTTGACAACTTCCTCTGTGGGGTACCATCCTCATCTCGGTGGGATGTCAG 402
QY 1162 TTTTCTGTGCCCTTATCGTCTGGTCTTTGTATGTCCAGGATGAAGAGAAAAATTGAAC 1221
DB 401 TTTTCTGTGCCCTTATCGTCTGGTCTTTGTATGTCCAGGATGAAGAGAAAAATTGAAC 342
QY 1222 GAGAAATAAGTGTAGTCTTCTGAAAGCCCTTAAATGAAAAAAGAAATAGCTTGAAG 1281
DB 341 GAGAAATAAGTGTAGTCTTCTGAAAGCCCTTAAATGAAAAAAGAAATAGCTTGAAG 282
QY 1282 AAGACCATGAAGAAACAAAGTTGTCTGTGGTGATATTGAAACAAAGCATCTGTTCTG 1341
DB 281 AAGACCATGAAGAAACAAAGTTGTCTGTGGTGATATTGAAACAAAGCATCTGTTCTG 222
QY 1342 AGGTAGGGCTGCCACTGTGCCCTCCAGGCTGTGGTGAGGAGAGAACAGTCTCATTTCA 1401
DB 221 AGGTAGGGCTGCCACTGTGCCCTCCAGGCTGTGGTGAGGAGAGAACAGTCTCATTTCA 162
QY 1402 AACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAGGCTTCCAGGCTGGACTTGAAG 1461
DB 161 AACTTGGAGATTTGGAGGAAGCTCC--AGAGAGAGAGGCTTCCAGGCTGGACTTGAAG 104
QY 1462 AGGAAACACATAGATAGCAGCGTGAATGGTGCAGTTCAGTTCCTTAATGGAACTTG 1521
DB 103 AGGAAACACATAGATAGCAGCGTGAATGGTGCAGTTCAGTTCCTTAATGGAACTTG 44
QY 1522 TCCAGTTTCACTCAAGCCGTGAGCAACCAAAATAAATCCAGTGG 1564
DB 43 TCCAGTTTCACTCAAGCCGTGAGCAACCAAAATAAATCCAGTGG 1

RESULT 10
US-09-867-701-2808
; Sequence 2808, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2808
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-867-701-2808

Query Match 13.1%; Score 430; DB 10; Length 441;
Best Local Similarity 99.8%; Pred. No. 5.8e-115;
Matches 441; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2830 TAAACATGAGAGCGTTTGACAGAGCATGCTCGGTTGTTGGTTTACACAGCTTCG 2889
DB 1 TAAACATGAGAGCGTTTGACAGAGCATGCTCGGTTGTTGGTTTACACAGCTTCG 60
QY 2890 CCTCATATGCACAGGATTTAAACAAACAAATATAACTACAACTTCCTTGTAGTCTCT 2949
DB 61 CCTCATATGCACAGGATTTAAACAAACAAATATAACTACAACTTCCTTGTAGTCTCT 120
QY 2950 TATATAAGTAGAGTCTTGGTACTCTGCCCTCCTGTCAGTAGTGGCAGGATCTATTGGCA 3009
DB 121 TATATAAGTAGAGTCTTGGTACTCTGCCCTCCTGTCAGTAGTGGCAGGATCTATTGGCA 180
QY 3010 TATTCGGGAGCTTCTTAGAGGATGAGTTCTTTGAACACAGTGAAATTTAAATTAGTA 3069
DB 181 TATTCGGGAGCTTCTTAGAGGATGAGTTCTTTGAACACAGTGAAATTTAAATTAGTA 240
QY 3070 ACTTTTTTGCAGCAGTTTATTGACTGTTATTGCTAAGAGAGTAGTAAAGAAAGAAAGCC 3129
DB 241 ACTTTTTTGCAGCAGTTTATTGACTGTTATTGCTAAGAGAGTAGTAAAGAAAGAAAGCC 300
QY 3130 TGTGGCAATCTTGGTTATTTCTTTTAAGATTTCTGGCAGTGTGGGATGGATGAATGAAGT 3189
DB 301 TGTGG- AATCTTGGTTATTTCTTTTAAGATTTCTGGCAGTGTGGGATGGATGAATGAAGT 359
QY 3190 GGAATGTGAACTTTGGGCAAGTTAAATGGACAGCTTTCATGTTCTATTGTTCTACTCT 3249
DB 360 GGAATGTGAACTTTGGGCAAGTTAAATGGACAGCTTTCATGTTCTATTGTTCTACTCT 419
QY 3250 TAACTGAATAAAGAAAGCCCTACA 3271
DB 420 TAACTGAATAAAGAAAGCCCTACA 441

RESULT 11

US-09-867-701-3261/c
; Sequence 3261, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3261
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-3261

Query Match 12.4%; Score 406.4; DB 10; Length 414;
Best Local Similarity 99.8%; Pred. No. 4.4e-108;
Matches 407; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2872 TGTTTTTCCAGCTTCTGCCCTCACATGCACAGGATTTAAACAAACAAATATAACTACA 2931
DB 408 TGTTTTTCCAGCTTCTGCCCTCACATGCACAGGATTTAAACAAACAAATATAACTACA 349
QY 2932 ACTTCCCTTGTAGTCTTATATAAGTAGAGTCTTGGTACTCTGCCCTCCTGTCAGTAG 2991
DB 348 ACTTCCCTTGTAGTCTTATATAAGTAGAGTCTTGGTACTCTGCCCTCCTGTCAGTAG 289
QY 2992 TGGCAGGATCTATTGGCATATTCGGAGCTTCTTAGAGGATGAGGTTCTTTTGAACACAG 3051

DB 288 TGGCAGGATCTATTGGCATATTCGGAGCTTCTTAGAGGATGAGGTTCTTTTGAACACAG 229
QY 3052 TGAATAATTTAAATTTAGTAACTTTTTTGGCAAGCAGTTTATTGACTGTTATTGCTTAAGAAGA 3111
DB 228 TGAATAATTTAAATTTAGTAACTTTTTTGGCAAGCAGTTTATTGACTGTTATTGCTTAAGAAGA 169
QY 3112 AGTAAGAAAGAAAGAAAGCCCTGTTGGCAATCTTGGTATTCTTTTAAAGATTCTGGCAGTGT 3171
DB 168 AGTAAGAAAGAAAGAAAGCCCTGTTGGCAATCTTGGTATTCTTTTAAAGATTCTGGCAGTGT 109
QY 3172 GGGATGGATGAATGAAGTGGATGAACTTTGGCAAGTAAATGGGACAGCTTCCAT 3231
DB 108 GGGATGGATGAATGAAGTGGATGAACTTTGGCAAGTAAATGGGACAGCTTCCAT 49
QY 3232 GTTCATTCTGCTACCTCTTTAACTGAATAAAGAAAGCCCTACAGTTTTTTAG 3279
DB 48 GTTCATTCTGCTACCTCTTTAACTGAATAAAGAAAGCCCTACAGTTTTTTAG 1

RESULT 12

US-09-867-701-2909/c
; Sequence 2909, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2909
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-2909

Query Match 12.0%; Score 394.4; DB 10; Length 396;
Best Local Similarity 99.7%; Pred. No. 1.4e-104;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2886 TCTGCCCTCACATGCACAGGATTTAAACAAACAAATATAACTACAACTTCCCTTGTAGT 2945
DB 396 TCTGCCCTCACATGCACAGGATTTAAACAAACAAATATAACTACAACTTCCCTTGTAGT 337
QY 2946 CTCCTTATATAAGTAGAGTCTTGGTACTCTGCCCTCCTGTCAGTAGTGGCAGGATCTATT 3005
DB 336 CTCCTTATATAAGTAGAGTCTTGGTACTCTGCCCTCCTGTCAGTAGTGGCAGGATCTATT 277
QY 3006 GGCATATTCGGGAGCTTCTTAGAGGATGAGTCTTTTGAACACAGTGAAATTTAAATTT 3065
DB 276 GGCATATTCGGGAGCTTCTTAGAGGATGAGTCTTTTGAACACAGTGAAATTTAAATTT 217
QY 3066 AGTAACTTTTTTGAAGCAGGATTTATTGACTGTTATTGCTTAAGAAAGAAAGAAAGAA 3125
DB 216 AGTAACTTTTTTGAAGCAGGATTTATTGACTGTTATTGCTTAAGAAAGAAAGAAAGAA 157
QY 3126 AGCCTGTTGGCAATCTTGGTATTCTTTTAAAGATTCTGGCAGTGTGGGATGGATGAATG 3185
DB 156 AGCCTGTTGGCAATCTTGGTATTCTTTTAAAGATTCTGGCAGTGTGGGATGGATGAATG 97
QY 3186 AAGTGAATGTGAATCTTTGGGCAAGTTAAATGGGACAGCTTCCATGTTTCATTTGTCTAC 3245
DB 96 AAGTGAATGTGAATCTTTGGGCAAGTTAAATGGGACAGCTTCCATGTTTCATTTGTCTAC 37
QY 3246 CTCCTTAACTGAATAAAGAAAGCCCTACAGTTTTTTAGAA 3281
DB 36 CTCCTTAACTGAATAAAGAAAGCCCTACAGTTTTTTAGAA 1

Query Match 11.9%; Score 393; DB 10; Length 643;
Best Local Similarity 91.7%; Pred. No. 5.2e-104;
Matches 465; Conservative 0; Mismatches 35; Indels 7; Gaps 5;

Qy	1805	CTCCAGAAGCGAATTTCGAATCGACAGCTTACACCACTTACTCGAATTCGTGTCTGACCT	1864
Dp	505	CTCCAGAAGCGAATTTCGAATCGACAGCTTACACCACTTACTCGAATTCGTGTCTGACCT	448

1924 AGGAAAGTAAATGGCTCT-CTTAGAAGAAATGGTATGACCAGGATAAG-CCTGAAGTCTCTC 1980

QY

Dh 187 AGAAAAGTAAAGGTTCTTCTAGAAAAGAAATGGTATGCCCAGGATAAGCCTGAAGTCTCTC 328

Qy	1981	TCC - TCCTCCAGTTCG	CAGATCCTT	PACGCTGCTTTGGGTCA	TTGCCCATGTTGGC	2039
ph	127	TCGCTCTTTCCG	TCCTGTCAGT	TCTTTCCTTTCG	TCATTCGCTTCATTCGCTTCATTCGCTTC	268

[illegible]

QY	2100	GATGTTTCTTCAAAAGTGGCAACACCAATATGGCTTCTACTCTATGTTGGGTGGTATC	2159
pH	207	GATGTTTCTTCAAAAGTGGCAACACCAATATGGCTTCTACTCTATGTTGGGTGGTATC	148

[illegible][illegible][illegible]

RESULT 15

; Sequence 9834, Application US/09867701
 ; Patent No. US20020132237A1
 ; GENERAL INFORMATION:

; APPLICANT: Jones, Robert
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.497
 ; CURRENT APPLICATION NUMBER: US/09/867,701
 ; CURRENT FILING DATE: 2001-05-29

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9834
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9834

Query Match 9.9%; Score 324.4; DB 10; Length 338;
Best Local Similarity 99.4%; Pred. No. 3.7e-84;
Matches 336; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 2941 GTAGTCTCTTATATAAGTAGAGTCCTTGGTACTCT-GCCCTCCTGTGTCAGTAGTGGCAGGA 2999
Db |||||||
Qy 338 GTAGTCTCTTATATAAGTAGAGTCCTTGGTACTCTGCCCCCTCCTGTGTCAGTAGTGGCAGGA 279
Db |||||||
Qy 3000 TCTATTGGGCATATTCCGGAGCTTCTTAGAGGGATGAGGTTCTTTGAACACAGTCAAAATT 3059
Db |||||||
Qy 278 TCTATTGGGCATATTCCGGAGCTTCTTAGAGGGATGAGGTTCTTTGAACACAGTCAAAATT 219
Db |||||||
Qy 3060 TAAATTAGTAACTTTTGGCAAGCAGTTTATTGACTGTTATTGCTAAGAAGAGTAAGAA 3119
Db |||||||
Qy 218 TAAATTAGTAACTTTTGGCAAGCAGTTTATTGACTGTTATTGCTAAGAAGAGTAAGAA 159
Db |||||||
Qy 3120 AGAAAAGCCTGTGGCAATCTTGGTTATTCTTTAGATTCTTGGCAGTGTGGATGGA 3179
Db |||||||
Qy 158 AGAAAAGCCTGTGGCAATCTTGGTTATTCTTTAAGATTCTTGGCAGTGTGGATGGA 99
Db |||||||
Qy 3180 TGAATGAAGTGGAAATGTGAACCTTGGGCAAGTTAAATGGGACAGCCCTTCCATGTTCAATT 3239
Db |||||||
Qy 98 TGAATGAAGTGGAAATGTGAACCTTGGGCAAGTTAAGTGGGACAGCCCTTCCATGTTCAATT 39
Db |||||||
Qy 3240 GTCTACCTCTTAAGTGAATAAAAAGCCTACAGTTTTT 3277
Db |||||||
Qy 38 GTCTACCTCTTAAGTGAATAAAAAGCCTACAGTTTTT 1
Db |||||||

Search completed: June 28, 2003, 12:28:58
Job time : 317.679 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2003, 02:12:40 ; Search time 3041.91 Seconds
(without alignments)
17516.326 Million cell updates/sec

Title: US-09-981-353-172

Perfect score: 3290

Sequence: 1 cggcctctctcgtggttc.....agtttttagaaaaaaa 3290

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2246.4	68.3	3720	11	BC016752
2	948.8	28.8	950	9	AL544984
C 3	936.4	28.5	1017	9	AL581504
C 4	927.2	28.2	979	9	AL570991
5	925.8	28.1	1013	9	AL559985
6	911.8	27.7	977	14	BQ671812

7	904.4	27.5	995	14	BQ671754
8	893.6	27.2	978	14	BQ33019
9	876.8	26.7	996	14	BM807130
10	875.6	26.6	999	14	BQ669841
11	846.6	25.7	952	14	BQ687887
12	825.8	25.1	896	14	BQ423949
13	811.4	24.7	952	14	BQ670753
14	800	24.3	1069	14	BM906416
15	790.6	24.0	938	14	BQ921653
16	790.2	24.0	880	14	BQ224397
17	778.8	23.7	1051	13	BM455409
18	778.2	23.7	905	14	BQ684572
19	759.6	23.1	887	12	BQ477821
20	758.4	23.1	1068	13	BM559000
21	756.8	23.0	915	14	BQ224706
22	752.6	23.0	799	9	AU132849
23	752.6	22.9	908	13	BI869902
24	751.4	22.8	1022	14	BQ049281
25	747.6	22.7	892	14	BQ213616
26	742.4	22.6	1063	13	BM544407
27	737	22.4	896	14	BQ687655
28	735	22.3	841	9	AU131432
29	734.8	22.3	848	14	BQ670075
30	733.4	22.3	775	9	AU133064
31	732.8	22.3	748	13	BI254485
32	720.2	21.9	977	12	BQ699340
33	717.4	21.8	840	14	BQ437280
34	717.2	21.8	763	13	BI546428
35	713.4	21.7	752	9	AU141833
36	711.2	21.6	741	9	AU132418
37	709.6	21.6	812	13	BI199063
38	709.2	21.6	993	14	BQ685992
C 39	707.6	21.5	748	14	BQ575524
C 40	695.6	21.1	720	14	BQ182587
41	695.6	21.1	763	13	BI545698
42	693	21.1	733	13	BI752094
43	692.4	21.0	980	14	BQ893809
C 44	692	21.0	697	14	BQ014211
45	691.2	21.0	915	14	BQ713677

ALIGNMENTS

RESULT 1	BC016752	3720 bp	mRNA	linear	HTC 09-NOV-2001
LOCUS	BC016752				
DEFINITION	Homo sapiens, clone IMAGE:2959721, mRNA.				
ACCESSION	BC016752				
VERSION	BC016752.1	GI:16876952			
KEYWORDS	HTC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3720)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting				

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 10 Row: c Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4097194
This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

source
1..3720
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2959721"
/cissue type="Colon, adenocarcinoma"
/clone_lib="NIH MGC 15"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
BASE COUNT 1042 a 714 c 840 g 1124 t
ORIGIN

Query Match 68.3%; Score 2246.4; DB 11; Length 3720;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2264; Conservative 0; Mismatches 1; Indels 7; Gaps 1;

QY	1019	TAATGGTTTCGAGCTTTGCGAGCTTTTCTATGCTGCGACAGTTTGAATAAACCTCTTTTC	1078
DB	10	TAATGGTTTCGAGCTTTGCGAGCTTTTCTATGCTGCGACAGTTTGAATAAACCTCTTTTC	69
QY	1079	CATCATGTATACATGGAGACCGTTGCGCTTTGCAAACTTCCTCTGTGGGGTACCAT	1138
DB	70	CATCATGTATACATGGAGACCGTTGCGCTTTGCAAACTTCCTCTGTGGGGTACCAT	129
QY	1139	CCTCATCTCGTGGGATGTCAGTTTCTGTGCGCTTATCGTCTGTGTTCTTTCTATGTGC	1198
DB	130	CCTCATCTCGTGGGATGTCAGTTTCTGTGCGCTTATCGTCTGTGTTCTTTCTATGTGC	182
QY	1199	CAGGATCAAGAGAAAAATGAACGAGAAAAATAAGTGTAGTCTTCTGAAAGCCCCCTTAAT	1258
DB	183	CAGGATCAAGAGAAAAATGAACGAGAAAAATAAGTGTAGTCTTCTGAAAGCCCCCTTAAT	242
QY	1259	GGAAAAAAGAAATAGCTTGAAGAGAACCATGAAGAAACAAGTTGTCTGTGGTATAT	1318
DB	243	GGAAAAAAGAAATAGCTTGAAGAGAACCATGAAGAAACAAGTTGTCTGTGGTATAT	302
QY	1319	TGAAACAAGCATCTCTGTTTCTCAGGTAGGCGCTGCCACTGCGCCCTCCAGGCTGTGGT	1378
DB	303	TGAAACAAGCATCTCTGTTTCTCAGGTAGGCGCTGCCACTGCGCCCTCCAGGCTGTGGT	362
QY	1379	GGAGGAGAGAACAGTCTCTATTCAAACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAG	1438
DB	363	GGAGGAGAGAACAGTCTCTATTCAAACTTGGAGATTTGGAGGAAGCTCCAGAGAGAGAG	422
QY	1439	GCTTCCAGCGTGGAATTGAAGAGGAAACAGCATAGATAGCAGCGTGAATGGTGCAGT	1498
DB	423	GCTTCCAGCGTGGAATTGAAGAGGAAACAGCATAGATAGCAGCGTGAATGGTGCAGT	482
QY	1499	GCAGTTGCCCTAATGGGAACCTTGTCCAGTTTCAGTCAAGCCGTGAGCAACCAATAAATC	1558
DB	483	GCAGTTGCCCTAATGGGAACCTTGTCCAGTTTCAGTCAAGCCGTGAGCAACCAATAAATC	542
QY	1559	CAGTGGCCACTACCAAGTATCACACCGTGCATAAGGATTCGGCGCTGTACAAAGAGCTACT	1618
DB	543	CAGTGGCCACTACCAAGTATCACACCGTGCATAAGGATTCGGCGCTGTACAAAGAGCTACT	602
QY	1619	CCATAATATCATCTTCCAGGTGGAGATTCGATGGGAGACTCCGGTGACAAACCTT	1678
DB	603	CCATAATATCATCTTCCAGGTGGGAGATTCGATGGGAGACTCCGGTGACAAACCTT	662
QY	1679	AAGCGCAATAAGCTATATCTTATACCATGGCAATATGGGATGCTCTGGATTC	1738
DB	663	AAGCGCAATAAGCTATATCTTATACCATGGCAATATGGGATGCTCTGGATTC	722
QY	1739	ATTCCTGCGCAAGAGGTGAACAGAGGCGCAAGAAATGGAGAAGCTGACATGGGCTAA	1798

DB	723	ATTCGCTGCGCAAGAGGTGAACAGAGGGCGAAGAAATGGAGAGCTGACATGGCCTAA	782
QY	1799	TGCAGACTCCAAAGAGCGAAATTCGAATGGACAGTTTACACAGTTACTGCAATGCTGTGTC	1858
DB	783	TGCAGACTCCAAAGAGCGAAATTCGAATGGACAGTTTACACAGTTACTGCAATGCTGTGTC	842
QY	1859	TGACCTTCCACTCAGCATCTGAGATAGACATAGTGTCAAGGCGAGAGATGGGTCTAGGTGA	1918
DB	843	TGACCTTCCACTCAGCATCTGAGATAGACATAGTGTCAAGGCGAGAGATGGGTCTAGGTGA	902
QY	1919	CAGAAAAAGGAAGTAATGGCTCTCTAGAGAAATGGTATGACAGGATAGCCTGAAGTCTTC	1978
DB	903	CAGAAAAAGGAAGTAATGGCTCTCTAGAGAAATGGTATGACAGGATAGCCTGAAGTCTTC	962
QY	1979	TCTCCTCTTCCAGTTCTCTGAGATCCTTACAGCCTGCTTGGGTGATTCGCCCATGGTGG	2038
DB	963	TCTCCTCTTCCAGTTCTCTGAGATCCTTACAGCCTGCTTGGGTGATTCGCCCATGGTGG	1022
QY	2039	CAATGACGTAAGCAATGGCCTCTGGCTTCTTATATTTGGTTTATGACACAGG	2098
DB	1023	CAATGACGTAAGCAATGGCCTCTGGCTTCTTATATTTGGTTTATGACACAGG	1082
QY	2099	AGATGTTTCTTCAAAAGTGGCAACCAATATGGCTTCTATCTATGTTGGTGGTAT	2158
DB	1083	AGATGTTTCTTCAAAAGTGGCAACCAATATGGCTTCTATCTATGTTGGTGGTAT	1142
QY	2159	CTGTCTGGTCTGTGGTTTGGGAGAGAGTATCCAGACCATGGGGAAGGATCTGAC	2218
DB	1143	CTGTCTGGTCTGTGGTTTGGGAGAGAGTATCCAGACCATGGGGAAGGATCTGAC	1202
QY	2219	ACCGATCACACCTCTAGTGGCTTCAGTATTGAATGGCATCTGCCCTCACTGTGGTAT	2278
DB	1203	ACCGATCACACCTCTAGTGGCTTCAGTATTGAATGGCATCTGCCCTCACTGTGGTAT	1262
QY	2279	TGATCAAAATATTTGGCTTCCCATCAGTACAAACATTTGAAGTGGGCTCTGTGTGTC	2338
DB	1263	TGATCAAAATATTTGGCTTCCCATCAGTACAAACATTTGAAGTGGGCTCTGTGTGTC	1322
QY	2339	TGTTGGCTGGCTCCGCTCCAGAGGCTGTTGACGTGGGCTCTTTTCGTAACATTTTAT	1382
DB	1323	TGTTGGCTGGCTCCGCTCCAGAGGCTGTTGACGTGGGCTCTTTTCGTAACATTTTAT	1382
QY	2399	GGCCTGGTTTGTACAGTCCCCATTTCTGGAGTTATCAGTGTGCCATCATGCCAATCTT	2458
DB	1383	GGCCTGGTTTGTACAGTCCCCATTTCTGGAGTTATCAGTGTGCCATCATGCCAATCTT	1442
QY	2459	CAGATATGTCATCTCAGAAATGTGAAGCTGTTTGAAGTTAAATTTGTCTCAATGTTGG	2518
DB	1443	CAGATATGTCATCTCAGAAATGTGAAGCTGTTTGAAGTTAAATTTGTCTCAATGTTGG	1502
QY	2519	GACCATCTTAGCTATTTCTGCTCCCTGAGAAATGATTTACAGTGTAAACAGAGAGCTGAC	2578
DB	1503	GACCATCTTAGCTATTTCTGCTCCCTGAGAAATGATTTACAGTGTAAACAGAGAGCTGAC	1562
QY	2579	AAGAGTCTTTTATTTGGAGCCAGAGGAGGAGTGTACTTGTGCTATAAATGCTGCTTT	2638
DB	1563	AAGAGTCTTTTATTTGGAGCCAGAGGAGGAGTGTACTTGTGCTATAAATGCTGCTTT	1622
QY	2639	GTGCTAAATATGAAATTTAGCTGTGTAATAATAGCCGGGTTCCACTGGCTC	2698
DB	1623	GTGCTAAATATGAAATTTAGCTGTGTAATAATAGCCGGGTTCCACTGGCTC	1682
QY	2699	CTGCTGAGGCTCCCTTCTCTGGGCTGTGAATTCCTGACATATTTCTACTTTTTG	2758
DB	1683	CTGCTGAGGCTCCCTTCTCTGGGCTGTGAATTCCTGACATATTTCTACTTTTTG	1742
QY	2759	TATCAGGCTTCAATTTCCATTATGTTTAAATGTTGCTCTGAAAGTGAATTTGATTTTTT	2818
DB	1743	TATCAGGCTTCAATTTCCATTATGTTTAAATGTTGCTCTGAAAGTGAATTTGATTTTTT	1802
QY	2819	TTTCTTTTTTTTAAACCATGAAGCGGTTTGAAGAGATGCTCTGCGTGTGTTGTTTC	2878
DB	1803	TTTCTTTTTTTTAAACCATGAAGCGGTTTGAAGAGATGCTCTGCGTGTGTTGTTTC	1862

```
QY 2879 ACCAGCTTCGCTCCATGCGACAGGATTTTAAACAAAAATATAAATACAACTTCCC 2938
D 1863 ACCAGCTTCGCTCCATGCGACAGGATTTTAAACAAAAATATAAATACAACTTCCC 1922
QY 2939 TTGAGTCTCTTATTAAGTAGAGTCTTGGTACTCTGCCCTCTGTGAGTAGTGGCAGG 2998
D 1923 TTGAGTCTCTTATTAAGTAGAGTCTTGGTACTCTGCCCTCTGTGAGTAGTGGCAGG 1982
QY 2999 ATCTATTGGCATATTCGGGAGCTTTCTAGAGGATGAGGTTCTTTGAAACACAGTGAAT 3058
D 1983 ATCTATTGGCATATTCGGGAGCTTTCTAGAGGATGAGGTTCTTTGAAACACAGTGAAT 2042
QY 3059 TTAATATTAGTAATCTTTTTCGAAGCAGTATTATTTGACTGTTTATTTGCTTAAGAAGTAAGA 3118
D 2043 TTAATATTAGTAATCTTTTTCGAAGCAGTATTATTTGACTGTTTATTTGCTTAAGAAGTAAGA 2102
QY 3119 AAAAAAAGCCTCTTGGCAATCTTGGTATTTCTTTAAAGATTTCTGGCAGTGGGATGG 3178
D 2103 AAAAAAAGCCTCTTGGCAATCTTGGTATTTCTTTAAAGATTTCTGGCAGTGGGATGG 2162
QY 3179 ATGAATGAAGTGAATGTGAATCTTGGCAAGTTAAATGGGACAGCTTCCATGTTCAAT 3238
D 2163 ATGAATGAAGTGAATGTGAATCTTGGCAAGTTAAATGGGACAGCTTCCATGTTCAAT 2222
QY 3239 TGCTACCTCTTAATGAATATAAAGCCTTACAGTATTTTAGAAAAA 3290
D 2223 TGCTACCTCTTAATGAATATAAAGCCTTACAGTATTTTAGAAAAA 2274
```

RESULT 2

```
AL544984 950 bp mRNA linear EST 16-FEB-2001
LOCUS AL544984 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1013Y03 5
DEFINITION prime, mRNA sequence.
ACCESSION AL544984
VERSION AL544984.1 GI:12877465
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 950)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
```

FEATURES

```
source
1..950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1013Y03"
/clone_lib="LTI_NFL006_P12"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
```

BASE COUNT

249 a 211 c 240 g 247 t 3 others

ORIGIN

Query Match 28.8%; Score 948.8; DB 9; Length 950;
Best Local Similarity 99.7%; Pred. No. 4.7e-200;
Matches 947; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
QY 1452 GACTTTGAAAGAGAAACCAAGCATAGATAGCACCGTGAATGGTGCAGTGCAGTGCCTAAT 1511
D 1 GACTTTGAAAGAGAAACCAAGCATAGATAGCACCGTGAATGGTGCAGTGCAGTGCCTAAT 60
QY 1512 GGGAACTTTGTCCAGTTCAAGTCAAGCCGTGAGCAACCAATAAATCTCAGTGGCCACTAC 1571
D 61 GGGAACTTTGTCCAGTTCAAGTCAAGCCGTGAGCAACCAATAAATCTCAGTGGCCACTAC 120
QY 1572 CAGTATCACACCGTGCATTAAGGATTCGGCCCTGTACAAAGAGCTACTCCATAAATTACAT 1631
D 121 CAGTATCACACCGTGCATTAAGGATTCGGCCCTGTACAAAGAGCTACTCCATAAATTACAT 180
QY 1632 CTTGCCAAGGTGGGAGATTGTCATGGGAGACTCCGGTGACAAACCCCTTAAGGGGCAATAAT 1691
D 181 CTTGCCAAGGTGGGAGATTGTCATGGGAGACTCCGGTGACAAACCCCTTAAGGGGCAATAAT 240
QY 1692 AGCTATATCTCTATACCATATGTGGCATATGTGGCATGCCTCTGGATTCATTCGTGGCAAA 1751
D 241 AGCTATATCTCTATACCATATGTGGCATATGTGGCATGCCTCTGGATTCATTCGTGGCAAA 300
QY 1752 GAAGGTGAACAGAGAGGGCGAAGAAATGGAGAGCTGACATGCGCTAATGCAGACTCCAAG 1811
D 301 GAAGGTGAACAGAGAGGGCGAAGAAATGGAGAGCTGACATGCGCTAATGCAGACTCCAAG 360
QY 1812 AAGCGAATTCGAATGGACAGTTTACACCAAGTTTACACCAAGTTTACCAATGCTGTGCTGACTCA 1871
D 361 AAGCGAATTCGAATGGACAGTTTACACCAAGTTTACCAATGCTGTGCTGACTCA 420
QY 1872 GCATCTGAGATAGACATGAGTGTCAAGGAGAGATGGGTCTAGGTGACAGAAAGAAAGT 1931
D 421 GCATCTGAGATAGACATGAGTGTCAAGGAGAGATGGGTCTAGGTGACAGAAAGAAAGT 480
QY 1932 AATGGCTCTCTAGAGAAGTGTATGACACAGGATGAGCTCTCTCTCTCTCTCTCTCTCCAG 1991
D 481 AATGGCTCTCTAGAGAAGTGTATGACACAGGATGAGCTCTCTCTCTCTCTCTCTCTCCAG 540
QY 1992 TTCCTGACATCCTTACAGCCTGCTTTTGGGTCTTTCGCCCATGGTGGAATGACGTAAGC 2051
D 541 TTCCTGACATCCTTACAGCCTGCTTTTGGGTCTTTCGCCCATGGTGGAATGACGTAAGC 600
QY 2052 AATGCCATTGGGCTCTGTTGCTTTATATTTGGTTATGACAGGAGATGTTTCTTCA 2111
D 601 AATGCCATTGGGCTCTGTTGCTTTATATTTGGTTATGACAGGAGATGTTTCTTCA 660
QY 2112 AAGTGGCAACACCAATATGCGCTTCTACTATGGTGGTGGTATCTGTGTGCTGCTG 2171
D 661 AAGTGGCAACACCAATATGCGCTTCTACTATGGTGGTGGTATCTGTGTGCTGCTG 720
QY 2172 TGGGTTTGGGGAAGAGATTTATCCAGACCATGGGGAAGATCTGCACCCGATCACACCC 2231
D 721 TGGGTTTGGGGAAGAGATTTATCCAGACCATGGGGAAGATCTGCACCCGATCACACCC 780
QY 2232 TCTAGTGGCTTCAAGTATTGAACTGGCATCTGCCCTCTACCTGTGGTGAATGCGATCAATATT 2291
D 781 TCTAGTGGCTTCAAGTATTGAACTGGCATCTGCCCTCTACCTGTGGTGAATGCGATCAATATT 840
QY 2292 GCGCTTCCCATCAGTACACACATTTGTAAGTGGGCTCTGTTGTGCTGTGCTGGCTC 2351
D 841 GCGCTTCCCATCAGTACACACATTTGTAAGTGGGCTCTGTTGTGCTGTGCTGGCTC 900
QY 2352 CGGTCCAAGAAGCGTGTGACTGGGCTCTCTTTTCGTAACATTTTATGCG 2401
D 901 CGGTCCAAGAAGCGTGTGACTGGGCTCTCTTTTCGTAACATTTTATGCG 950
```

RESULT 3

AL581504/c
LOCUS AL581504 LTI_FL011_BCI Homo sapiens cDNA clone CS0D004YN16 3 prime
DEFINITION mRNA sequence.
ACCESSION AL581504
VERSION AL581504.1 GI:12948569

1017 bp mRNA linear EST 16-FEB-2001


```
Query Match      28.2%; Score 927.2; DB 9; Length 979;
Best Local Similarity 97.7%; Pred. No. 2.9e-195;
Matches 946; Conservative 4; Mismatches 17; Indels 1; Gaps 1;

QY 2263 CCCTCACTGTTGGTGAATTCATCAAAATATTGGCGCTTCCCATCAGTACAAACACATTTGTAAG 2322
DB 978 CCCTCACTGTTGGTGAATTCATCAAAATATTGGCGCTTCCCATCAGTACAAACACATTTGTAAG 920

QY 2323 TGGGCTCTGTGTGTCTGTGGTGGCTCCGCTCCAGAGAGGCTGTGTGAGTGGCGCTCTCT 2382
DB 919 TGGGCTCTGTGTGTCTGTGGTGGCTCCGCTCCAGAGAGGCTGTGTGAGTGGCGCTCTCT 860

QY 2383 TTGCTAACATTTTATGCGCTGTGTGTACAGTCCCAATTTCTGGAGTTATCAGTGTCTG 2442
DB 859 TTGCTAACATTTTATGCGCTGTGTGTACAGTCCCAATTTCTGGAGTTATCAGTGTCTG 800

QY 2443 CCATCATGCAATCTTTCAGATATGTCATCTCAGATGTGAAGCTCTTTGAGATTAAT 2502
DB 799 CCATCATGCAATCTTTCAGATATGTCATCTCAGATGTGAAGCTCTTTGAGATTAAT 740

QY 2503 TTGTGTCAATGTTTGGGACCATCTTAGTATTCTCTGCTCCCTCGAAGATGATTACAGTG 2562
DB 739 TTGTGTCAATGTTTGGGACCATCTTAGTATTCTCTGCTCCCTCGAAGATGATTACAGTG 680

QY 2563 TTAACAGAGACTGCAAGAAGTCTTTTATTTGGGAGCCAGAGAGGAGTGTACTTG 2622
DB 679 TTAACAGAGACTGCAAGAAGTCTTTTATTTGGGAGCCAGAGAGGAGTGTACTTG 620

QY 2623 TGCATTAACCTGCTTTTGTGCTAAATATGATTTCTCTCAAAATTAGTGTGTAATAAGCC 2682
DB 619 TGCATTAACCTGCTTTTGTGCTAAATATGATTTCTCTCAAAATTAGTGTGTAATAAGCC 560

QY 2683 CGGGTTCACATGCTCTCTGCTGAGGTCCTCTTCTCTGCGGTGTGAATTCCTGTACAT 2742
DB 559 CGGGTTCACATGCTCTCTGCTGAGGTCCTCTTCTCTGCGGTGTGAATTCCTGTACAT 500

QY 2743 ATTTCTCTACTTTTGTATCAGGCTTCAATTCATATATGTTTAAATTTGTCTCTCAAGA 2802
DB 499 ATTTCTCTACTTTTGTATCAGGCTTCAATTCATATATGTTTAAATTTGTCTCTCAAGA 440

QY 2803 TGACTTGTGATTTTCTTTTCTTTTAAACCATGAAGCCGTTTGAAGAGCATGCT 2862
DB 439 TGACTTGTGATTTTCTTTTCTTTTAAACCATGAAGCCGTTTGAAGAGCATGCT 380

QY 2863 CTGCGTGTGTGTTTCAACAGCTTCTGCCCTCACATGCACAGGGAATTAACACAAAAAT 2922
DB 379 CTGCGTGTGTGTTTCAACAGCTTCTGCCCTCACATGCACAGGGAATTAACACAAAAAT 320

QY 2923 ATAACACAACTTCCCTTGTAGTCTCTTATATAAGTAGAGTCTCTGCTACTCTGCCCTCC 2982
DB 319 ATAACACAACTTCCCTTGTAGTCTCTTATATAAGTAGAGTCTCTGCTACTCTGCCCTCC 260

QY 2983 TGTCACTAGTGGCAGATCTATTGGCATATTCCGGAGCTTCTTAGAGGAGTAGGTTCTT 3042
DB 259 TGTCACTAGTGGCAGATCTATTGGCATATTCCGGAGCTTCTTAGAGGAGTAGGTTCTT 200

QY 3043 TGAACACAGTGAATAATTAATTAAGTAACTTTTTCGAAGCAGTATTATGACGTATTG 3102
DB 199 TGAACACAGTGAATAATTAATTAAGTAACTTTTTCGAAGCAGTATTATGACGTATTG 140

QY 3103 CTAAGAAGAGTAGAAGAAAAAGGCTGTGGCAATCTTGGTATTCTTTTAAGATTTC 3162
DB 139 CTAAGAAGAGTAGAAGAAAAAGGCTGTGGCAATCTTGGTATTCTTTTAAGATTTC 80

QY 3163 TGGCAGTGGGATGAGTGAATGAAGTGAATGTGAATCTTTGGGCAAGTTAAATGGGACA 3222
DB 79 TGGCAGTGGGATGAGTGAATGAAGTGAATGTGAATCTTTGGGCAAGTTAAATGGGACA 20

QY 3223 GCCTTCCA 3230
DB 19 GCCTTCCA 12
```

```
RESULT 5
AL559985
LOCUS
DEFINITION
AL559985 LTI_FL011_BC1 Homo sapiens cDNA clone CS0DG004YN16 5 prime
mRNA sequence.
AL559985
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1013)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..1013
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DG004YN16"
/clone_lib="LTI_FL011_BC1"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
Library was constructed by Life Technologies. Contact:
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT 194 a 265 c 243 g 304 t 7 others
ORIGIN
Query Match 28.1%; Score 925.8; DB 9; Length 1013;
Best Local Similarity 96.3%; Pred. No. 5.9e-195;
Matches 972; Conservative 6; Mismatches 28; Indels 3; Gaps 3;
QY 7 TCTCTCGTGGTCTCTCTCTCGCGCTGAACCCCGCGCTCTCTCGGAGGTC 66
DB 1 TCTCTCGTGGTCTCTCTCTCGCGCTGAACCCCGCGCTCTCTCGGAGGTC 60
QY 67 GTGAGTCCGCTGAGCTGTCCCGGTGCGCGGCGGTGTGCCCGGTGTGCCGCTCCAG 126
DB 61 GTGAGTCCGCTGAGCTGTCCCGGTGCGCGGCGGTGTGCCCGGTGTGCCGCTCCAG 120
QY 127 CGGCTGCCGCTCGATCTCTCTCTCGCTCTCGGCTCCGCTCTCTCTCTCTCTCTCT 186
DB 121 CGGCTGCCGCTCGATCTCTCTCTCGCTCTCGGCTCCGCTCTCTCTCTCTCTCTCT 180
QY 187 GGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 246
DB 181 GGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 247 AGAACCACCCAGAACACAGATGCCATACGACGATAGCAGTAACTCCCGAGCTCGTT 306
DB 241 AGAACCACCCAGAACACAGATGCCATACGACGATAGCAGTAACTCCCGAGCTCGTT 300
QY 307 TCTGTGCCGTAGTTTACAGTATTTTATATATATATATATATATATATATATATATAT 366
DB 301 TCTGTGCCGTAGTTTACAGTATTTTATATATATATATATATATATATATATATATAT 360
QY 367 TTGATACCTCATATTTCTGTGTTTACATCTTGAAGCGCTCAGTAGTCTCTTCTTAAAC 426
DB 361 TTGATACCTCATATTTCTGTGTTTACATCTTGAAGCGCTCAGTAGTCTCTTCTTAAAC 420
```

QY 427 AACCACTACTCCAGAGAAATGGCAACGCTGATTACCACTACTACAGCTGCTACCGCGCTT 486
 Db 421 AACCACTACTCCAGAGAAATGGCAACGCTGATTACCACTACTACAGCTGCTACCGCGCTT 480
 QY 487 CTGGTCTTTGGTGGACTACTATGGATGCTCATCTCGGGCTTCATTATTGCAATTTGTCT 546
 Db 481 CTGGTCTTTGGTGGACTACTATGGATGCTCATCTCGGGCTTCATTATTGCAATTTGTCT 540
 QY 547 TGGCATTCTCCGTGGAGGCAATGATGTAGCAAAATCTTTTGTGTAGCTGTGGGCTCAG 606
 Db 541 TGGCATTCTCCGTGGAGGCAATGATGTAGCAAAATCTTTTGTGTAGCTGTGGGCTCAG 600
 QY 607 GTCTAGTGACCTGAAGCAAGCTGCTATCTAGCTAGCATCTTTTGAACAGTGGGCTCTG 666
 Db 601 GTCTAGTGACCTGAAGCAAGCTGCTATCTAGCTAGCATCTTTTGAACAGTGGGCTCTG 660
 QY 667 TCTTACTGGGGGCCAAAGTGAAGCAACCATCCGGAAGGGCTTGAATGACGTGGAGATGT 726
 Db 661 TCTTACTGGGGGCCAAAGTGAAGCAACCATCCGGAAGGGCTTGAATGACGTGGAGATGT 720
 QY 727 ACAATCTGACTAAGGGCTACTGATGCGGGCTCAGTCAGTGTCTATGTTTGGTTCTGCTG 786
 Db 721 ACAATCTGACTAAGGGCTACTGATGCGGGCTCAGTCAGTGTCTATGTTTGGTTCTGCTG 780
 QY 787 TGTGGCAACTCGTGGCTTCGTTTGAAGCTCCCTATTTCTGAAACCCATTTGATTTGTTG 846
 Db 781 TGTGGCAACTCGTGGCTTCGTTTGAAGCTCCCTATTTCTGAAACCCATTTGATTTGTTG 840
 QY 847 GTGCAACTATTGTTTCTCCCTCGTGGCAAGGGCGAGGGGTGCAAGTGTCTGAAAC 906
 Db 841 GTGC-ACTATTGTTTCTCCCTCGTGGCAAGGGCGAGGGGTGCAAGTGTCTGAAAC 899
 QY 907 TGATAAAATGTGATGCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
 Db 900 TGATAAAATGTGATGCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958
 QY 967 TTTTATTCTCTCGTTCGTCATCTCCATCAATGAGGAGATCAGTTCTTAAT 1015
 Db 959 -TTTATTCTCTCGTTCGTCATCTCCATCAATGAGGAGATCAGTTCTTAAT 1006

RESULT 6

BO671812 977 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT_8189344 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6255115
 DEFINITION 5', mRNA sequence.

ACCESSION BO671812
 VERSION BO671812.1 GI:21782646
 KEYWORDS EST.

SOURCE

human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 977)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

AUTHORS

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: ATCC

JOURNAL

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov

COMMENT

High quality sequence stop: 725.

FEATURES

Location/Qualifiers

source

1..977
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/clone="IMAGE:6255115"
 /clone_lib="NIH_MGC_102"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-CDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies). Note: this is a NIH MGC Library."
 283 a 216 c 240 g 236 t 2 others

BASE COUNT

Query Match 27.7%; Score 911.8; DB 14; Length 977;
 Best Local Similarity 97.9%; Pred. No. 7.5e-192;
 Matches 956; Conservative 0; Mismatches 17; Indels 4; Gaps 3;
 QY 1031 AGCTTTGCCAGTTTCTATGCTGCACAGTTGGAATAAACCTCTTTTCCCATCATGTATAC 1090
 Db 1 AGCTTTGCCAGTTTCTATGCTGCACAGTTGGAATAAACCTCTTTTCCCATCATGTATAC 60
 QY 1091 TGGAGCACCGTTGCTGGGCTTTGACAAACTTCTCTGTGGGTACCATCTCATCTCGGT 1150
 Db 61 TGGAGCACCGTTGCTGGGCTTTGACAAACTTCTCTGTGGGTACCATCTCATCTCGGT 120
 QY 1151 GGGATGTGCAGTTTCTGTGCCCTTATCGTCTGGTCTTTGTATGTCCAGATGAAGAG 1210
 Db 121 GGGATGTGCAGTTTCTGTGCCCTTATCGTCTGGTCTTTGTATGTCCAGATGAAGAG 180
 QY 1211 AAAAATGAACGAGAAATAAGTGTAGTCTCTGAAAGCCCTTAAATGGAATAAAGAA 1270
 Db 181 AAAAATGAACGAGAAATAAGTGTAGTCTCTGAAAGCCCTTAAATGGAATAAAGAA 240
 QY 1271 TAGCTTTGAAAGAGAACCATGAAGAAACAAAGTTGTCTGTGTGTATATTGAAACAAAGCA 1330
 Db 241 TAGCTTTGAAAGAGAACCATGAAGAAACAAAGTTGTCTGTGTGTATATTGAAACAAAGCA 300
 QY 1331 TCCTGTTTCTGAGGTAGGGCTGCGACCTGTGCGCCCTCCAGGCTGTGGGAGAGAGAAC 1390
 Db 301 TCCTGTTTCTGAGGTAGGGCTGCGACCTGTGCGCCCTCCAGGCTGTGGGAGAGAGAAC 360
 QY 1391 AGTCTCATTAACCTTTGGAGATTTCGAGGAAGCTCCAGAGAGAGAGAGCTTCCAGCGT 1450
 Db 361 AGTCTCATTAACCTTTGGAGATTTCGAGGAAGCTCCAGAGAGAGAGAGCTTCCAGCGT 420
 QY 1451 GGACTTTGAAAGAGAAACCCAGCATAGATAGCACCGTGAATGTGCGAGTGCAGTTGCTTAA 1510
 Db 421 GGACTTTGAAAGAGAAACCCAGCATAGATAGCACCGTGAATGTGCGAGTGCAGTTGCTTAA 480
 QY 1511 TGGGAACCTTGTCCAGTTTCAGTCAAGCGTCAAGCAACCAATAAATCCAGTGGCCACTA 1570
 Db 481 TGGGAACCTTGTCCAGTTTCAGTCAAGCGTCAAGCAACCAATAAATCCAGTGGCCACTA 540
 QY 1571 CCAGTATCACACCGTGCATAAGGATTTCGGGCTGTACAAAGAGCTACTCCATATAATTA 1630
 Db 541 CCAGTATCACACCGTGCATAAGGATTTCGGGCTGTACAAAGAGCTACTCCATATAATTA 600
 QY 1631 TCTTGCACAGGTGGGAGATTTCATGGGAGACTCCGGTGAACAAACCCCTTAAGCGCGCAATA 1690
 Db 601 TCTTGCACAGGTGGGAGATTTCATGGGAGACTCCGGTGAACAAACCCCTTAAGCGCGCAATA 660
 QY 1691 TAGCTATATCTCTATACCATGGCAATATGTGGCATGCTCTGGATTCTATTTCGTGCCAA 1750
 Db 661 TAGCTATATCTCTATACCATGGCAATATGTGGCATGCTCTGGATTCTATTTCGTGCCAA 720
 QY 1751 AGAAGGTGAACAGAGGGCGAAGAAATGGAGAGCTGACATGCGCTTAATGCGAGCTCCAA 1810
 Db 721 AGAAGGTGAACAGAGGGCGAAGAAATGGAGAGCTGACATGCGCTTAATGCGAGCTCCAA 780
 QY 1811 GAAGCGAATTTCGAATGGACAGTTTACACCACTTACT-GCAATGCTGTGTGACCTTCACT 1869

Db 781 GAAGCGAATTCGAATGCACAGATTACACCACTTACTGGCAATGCTGTGCTGACCTTCCT 840
 Qy 1870 CAGCATCTGAGATAGACATGAGTGTCAAGSCAGAG-ATGGGTCTAGGTGACAGAAAGGA 1928
 Db 841 CAGCATCTGAGATAGACATGAGTGTCAAGSCAGAGAAATGGTCTAGGTGACAGAAAGG 900
 Qy 1929 AGTAATGGCTCTCTAGAGAAATGGTATGACCA--GGATAGCCTGAAAGTCTCTCTCTCT 1986
 Db 901 AGTAATGGCTCTCTAGAGAAATGGTATGACCA--GGATAGCCTGAAAGTCTCTCTCTCT 960
 Qy 1987 TCCAGTCTCTGCGAGATC 2003
 Db 961 TTCCGTTCTGCAATC 977

RESULT 7

BQ671754
 LOCUS BQ671754 995 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8189125 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6255034
 5', mRNA sequence.

ACCESSION BQ671754

VERSION BQ671754.1

KEYWORDS GI:21782588

SOURCE EST.

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

JOURNAL NIH-MGC http://mgi.nci.nih.gov/

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LCM2405 row: h column: 11

High quality sequence stop: 617.

Location/Qualifiers

1..995

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6255034"

/clone_lib="NIH_MGC_102"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: ealivary gland; Vector: pOTB7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed

by Ling Hong in the laboratory of Gerald M. Rubin

(University of California, Berkeley) using ZAP-cDNA

synthesis kit (Stratagene) and Superscript II RT (Life

Technologies). Note: this is a NIH MGC Library."

BASE COUNT 285 a 220 c 246 g 243 t 1 others

ORIGIN

Query Match

Best Local Similarity 27.5%; Score 904.4; DB 14; Length 995;

Matches 964; Conservative 0; Mismatches 27; Indels 4; Gaps 4;

Qy 1031 AGCTTTGCCAGTTTCTATGCTGCGACAGTTGGAATAAACCTCTTTTCCATCATGATAC 1090

Db 1 AGCTTTGCCAGTTTCTATGCTGCGACAGTTGGAATAAACCTCTTTTCCATCATGATAC 60

Qy 1091 TGGAGCACCGTTCGCTGGGCTTTGACAACTTCTCTGTGGGTACCATCTCATCTCGGT 1150

Db 61 TGGAGCACCGTTCGCTGGGCTTTGACAACTTCTCTGTGGGTACCATCTCATCTCGGT 120

Qy 1151 GGGATGTGCGAGTTTCTGTGCCCCCTTATCGTCTGGTCTTTTGTATGTCCAGGATGAAG 1210
 Db 121 GGGATGTGCGAGTTTCTGTGCCCCCTTATCGTCTGGTCTTTTGTATGTCCAGGATGAAG 180
 Qy 1211 AAAAATTTGAACGAGAAATAAAGTGTAGTCTTCTGAAAGCCCCCTTAATCGAAAAAAGAA 1270
 Db 181 AAAAATTTGAACGAGAAATAAAGTGTAGTCTTCTGAAAGCCCCCTTAATCGAAAAAAGAA 240
 Qy 1271 TAGCTTTGAAAGAACCAATGAAGAAACAAAGTTGTCTGTTGGTGATATTGAAAAACAAGCA 1330
 Db 241 TAGCTTTGAAAGAACCAATGAAGAAACAAAGTTGTCTGTTGGTGATATTGAAAAACAAGCA 300
 Qy 1331 TCCTGTTTCTGAGGTAGGCGCTGCGACCTGTGCCCCCTCCAGGCTGTGGAGAGAGAAC 1390
 Db 301 TCCTGTTTCTGAGGTAGGCGCTGCGACCTGTGCCCCCTCCAGGCTGTGGAGAGAGAAC 360
 Qy 1391 AGTCTCATTTCAACCTTTGGAGATTTTGGAGAACTCCAGAGAGAGAGAGGCTTCCCAGCGT 1450
 Db 361 AGTCTCATTTCAACCTTTGGAGATTTTGGAGAACTCCAGAGAGAGAGAGGCTTCCCAGCGT 420
 Qy 1451 GGACTTTGAAAGAGGAAACCCAGCATAGATAGACCCGTGAATGGTGCAGTGCAGTTGCCTAA 1510
 Db 421 GGACTTTGAAAGAGGAAACCCAGCATAGATAGACCCGTGAATGGTGCAGTGCAGTTGCCTAA 480
 Qy 1511 TGGGAACCTTTGTCAGTTTCAGTCAAGCCGTGACCAACCAAAATAAACTCCAGTGGCCACTA 1570
 Db 481 TGGGAACCTTTGTCAGTTTCAGTCAAGCCGTGACCAACCAAAATAAACTCCAGTGGCCACTA 540
 Qy 1571 CCAGTATCACACCGTGCATAAGGATTCGGGCTGTACAAAGAGCTACTTCATTAATTTACA 1630
 Db 541 CCAGTATCACACCGTGCATAAGGATTCGGGCTGTACAAAGAGCTACTTCATTAATTTACA 600
 Qy 1631 TCTTGCCAA--GGTGGGAGATTGCATGGGAGACTCCGGTGACAAACCCCTTAAGGCGCAATA 1689
 Db 601 TCTTGCCAAAGGTGGGAGATTGCATGGGAGACTCCGGTGACAAACCCCTTAAGGCGCAATA 660
 Qy 1690 ATAGCTATATCTTCTATACCATGGCAATATGTGGCATGCTCTGGATTCATTCCGTGCCA 1749
 Db 661 ATAGCTATATCTTCTATACCATGGCAATATGTGGCATGCTCTGGATTCATTCCGTGCCA 720
 Qy 1750 AAGAAGGTGAACAGAAAGGCGAAGAAATGGAGAACTGACATGGCTTAATGCAGACTCCA 1809
 Db 721 AAGAAGGTGAACAGAAAGGCGAAGAAATGGAGAACTGACATGGCTTAATGCAGACTCCA 780
 Qy 1810 AGAAGCGAATTCGAATGGACACTTACCCAGTTACTGCAATCTGTGCTGACCTTCACT 1869
 Db 781 AGAAGCGAATTCGAATGGACACTTACCCAGTTACTGCAATCTGTGCTGACCTTCACT 840
 Qy 1870 CAGCATCTGAGATAGACATGATGTCAAGGCAG-AGATGGGTCTTAGGT-GACAGAAAAGG 1927
 Db 841 CCGCATCTGAAATAGACATGATGTCCAGGAGAAAAATGGGTCTTAGGTGGACGAAAAAGG 900
 Qy 1928 AAGTAATGGCTCTTAGAAGAAATGGTATGACCCAGATAAGCCTGAAGTC-TCTCTCTCT 1986
 Db 901 AAGTAATGGCTCTTAGAAGAAATGGTATGACCCAGATTAGCCTGAAATCTTCTCTCTCT 960
 Qy 1987 TCCAGTCTCTGCGAGATCTTACAGCCTCTCTTTGGG 2021
 Db 961 TTCCAGTCTCTGCGAAACCCCTAAACCCCTGTTTGG 995

RESULT 8

BQ933019

LOCUS BQ933019

DEFINITION AGENCOURT_8824319 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6459568

5', mRNA sequence.

ACCESSION BQ933019

VERSION BQ933019.1

KEYWORDS GI:22348402

SOURCE EST.

ORGANISM human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 978)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTP/Gazdar
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLC2644 row: n column: 17
 High quality sequence start: 10
 High quality sequence stop: 688.

FEATURES
 source
 1..978
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6459568"
 /clone_lib="NIH_MGC_18"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 277 a 213 c 258 g 230 t

ORIGIN
 Query Match 27.2%; Score 893.6; DB 14; Length 978;
 Best Local Similarity 97.8%; Pred. No. 8.3e-188;
 Matches 927; Conservative 0; Mismatches 19; Indels 2; Gaps 2;

QY 1097 ACCGTTGCTGGCTTTGACAACTCTCTGTGGGTACCATCTCATCTCGGTGGATG 1156
 DB 33 ACCGTTGCTGGCTTTGAC-AACTCTCTGTGGGTACCATCTCATCTCGGTGGATG 91
 QY 1157 TGCAGTTTCTGTGCCCTTATCGTCTGTGGTCTTTGTATGTCGCCAGATGAAGAGAAAAT 1216
 DB 92 TGCAGTTTCTGTGCCCTTATCGTCTGTGGTCTTTGTATGTCGCCAGATGAAGAGAAAAT 151
 QY 1217 TGAACGAGAAATAAAGTGTAGTCTCTTCTGAAAGCCCTTAAATGGAAGAAAAGAAATAGCTT 1276
 DB 152 TGAACGAGAAATAAAGTGTAGTCTCTTCTGAAAGCCCTTAAATGGAAGAAAAGAAATAGCTT 211
 QY 1277 GAAAGAGACCATGAAGAAACAAGTTGCTGTGTGATATTGAAACACAGCATCTGT 1336
 DB 212 GAAAGAGACCATGAAGAAACAAGTTGCTGTGTGATATTGAAACACAGCATCTGT 271
 QY 1337 TCTGAGTAGGCGCTGCCACTGTGCCCTTCCAGGCTGTGGTGAGAGAGACAGTCTC 1396
 DB 272 TCTGAGTAGGCGCTGCCACTGTGCCCTTCCAGGCTGTGGTGAGAGAGACAGTCTC 331
 QY 1397 ATTCAAACTTTGGAGATTGGAGGAAGCTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTT 1456
 DB 332 ATTCAAACTTTGGAGATTGGAGGAAGCTCCAGAGAGAGAGAGGCTTCCAGCGTGGACTT 391
 QY 1457 GAAAGAGAAACACAGATAGATAGCAGCGTGAATGGTGAGTGCAGTGCAGTTCCTAATGGAA 1516
 DB 392 GAAAGAGAAACACAGATAGATAGCAGCGTGAATGGTGAGTGCAGTGCAGTTCCTAATGGAA 451
 QY 1517 CTTGTGTCAGTTCAGTCAAGCGCTCAGCAACCAAAATAACTCCAGTGGCCACTACAGTA 1576
 DB 452 CTTGTGTCAGTTCAGTCAAGCGCTCAGCAACCAAAATAACTCCAGTGGCCACTACAGTA 511
 QY 1577 TCACACCGTGCATAAGGATTCCGGCTGTGTACAAAGAGCTACTCCATAAATTAACATCTTGC 1636

DB 512 TCACACCGTGCATAAGGATTCCGGCTGTGTACAAAGAGCTACTCCATAAATTAACATCTTGC 571
 QY 1637 CAAGGTGGGAGATTGCATGGGAGACTCCGGTCAACAAACCTTTAAGCGCGCAATAATAGCTA 1696
 DB 572 CAAGGTGGGAGATTGCATGGGAGACTCCGGTCAACAAACCTTTAAGCGCGCAATAATAGCTA 631
 QY 1697 TACTTCTTATACCATGGCAATATGTGGCAATGCTCTGGATTCATTCCTGCGCAAGAGG 1756
 DB 632 TACTTCTTATACCATGGCAATATGTGGCAATGCTCTGGATTCATTCCTGCGCAAGAGG 691
 QY 1757 TGAACAGAGGGCGAAGAAATGGAGAGCTGACATGGCTTAATGCAGACTCCCAAGAGG 1816
 DB 692 TGAACAGAGGGCGAAGAAATGGAGAGCTGACATGGCTTAATGCAGACTCCCAAGAGG 751
 QY 1817 AATTGCAATGGACAGTTACACCAAGTTACTGCAATGCTGTCTGACCTTCACTCAGCATC 1876
 DB 752 AATTGCAATGGACAGTTACACCAAGTTACTGCAATGCTGTCTGACCTTCACTCAGCATC 811
 QY 1877 TGAGATAGACATGAGTGTCAAGSCAGAGATGGTCTAGGTGACAGAAAGAGTAATGG 1936
 DB 812 TGAGATAGACATGAGTGTCAAGSCAGAGATGGTCTAGGTGACAGAAAGAGTAATGG 870
 QY 1937 CTCTCTAGAAATGATGATGACAGATAGGCTGAAGTCTCTCTCTTCCAGTTCT 1996
 DB 871 CTCTCTAGAAATGATGATGACAGATAGGCTGAAGTCTCTCTCTTCCAGTTCT 930
 QY 1997 CGAGATCTCTACAGCTGCTTTGGGTGCTTCGCCCATGGTGGCAATGA 2044
 DB 931 CGAGATCTCTAAGCTGCTTTGGGTGCTTCGCCCATGGTGGCAATGA 978

RESULT 9
 BM807130
 LOCUS BM807130
 DEFINITION AGENCOURT_6566577 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5738205
 5', mRNA sequence.
 ACCESSION BM807130
 VERSION BM807130.1 GI:19123953
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 996)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM12749 row: i column: 22
 High quality sequence stop: 745.
 Location/Qualifiers
 1..996
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5738205"
 /clone_lib="NIH_MGC_88"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: small intestine; Vector: pCMV-SPORT6;
 Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
 oligo-dT primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 221 a 227 c 234 g 314 t

ORIGIN
 source

Query Match	26.7%;	Score 876.8;	DB 14;	Length 996;
Best Local Similarity	99.1%;	Pred. No. 4.4e-184;		
Matches 892;	Conservative 0;	Mismatches 7;	Indels 1;	Gaps 1;
QY	245	AAAGAAACCCAGACAAACAGATGCGCCATACGAGCGGTATAGACGTAACTCCCGAGCTCGG	304	
DB	1	AACGAAACCCAGACAAACAGATGCGCTATACGCTGCGTATGCGAGTAACCTCCCGAGCTCGG	60	
QY	305	TTTCTGTGCGGTAGTTTACAGTATTTAAATTTATATATATATATATATATTTATTATATAGCAT	364	
DB	61	TTTCTGTGCGGTAGTTTACAGTATTTAAATTTATATATATATATATATATTTATTATATAGCAT	120	
QY	365	TTTGTAGTACCTCATATTTCTGTTTACACATCTTGAAAGCGCTCAGTAGTGTCTCTTACTAA	424	
DB	121	TTTGTAGTACCTCATATTTCTGTTTACACATCTTGAAAGCGCTCAGTAGTGTCTCTTACTAA	180	
QY	425	ACAACCACTACTCCAGAGAAATGGCAACGGTGTATTTACCACTACTACAGCTGCTACCGCGGC	484	
DB	181	ACAACCACTACTCCAGAGAAATGGCAACGGTGTATTTACCACTACTACAGCTGCTACCGCGGC	240	
QY	485	TTCTGTCTCTTTGGTGGACTACTATGATGCTCATCTCGGCTTCATTTATGCAATTTGT	544	
DB	241	TTCTGTCTCTTTGGTGGACTACTATGATGCTCATCTCGGCTTCATTTATGCAATTTGT	300	
QY	545	CTTGCGATTTCTCGGTGGAGCAATGATGTAGCAAAATTTCTTTGGTACAGCTGTGGGCTC	604	
DB	301	CTTGCGATTTCTCGGTGGAGCAATGATGTAGCAAAATTTCTTTGGTACAGCTGTGGGCTC	360	
QY	605	AGGTGTAGTGACCTCGAGCAAGCTGATCTAGTACAGTCTTTCGAAACAGTGGGCTC	664	
DB	361	AGGTGTAGTGACCTCGAGCAAGCTGATCTAGTACAGTCTTTCGAAACAGTGGGCTC	420	
QY	665	TGTCTTACTGGGGGCAAAAGTGGAGCAACACATCCGGAAGGCTTGATTGACGTGGAGAT	724	
DB	421	TGTCTTACTGGGGGCAAAAGTGGAGCAACACATCCGGAAGGCTTGATTGACGTGGAGAT	480	
QY	725	GTACAACTCGACTCAAGGCTCTACTGATGGCGGCTCAGTCACTAGTGTCTGTTGTTCTGC	784	
DB	481	GTACAACTCGACTCAAGGCTCTACTGATGGCGGCTCAGTCACTAGTGTCTGTTGTTCTGC	540	
QY	785	TGTGTGGCAACTCGTGGCTTCGTTTGAAGTCCCTATTTCTGGAACCCATTTGATTGT	844	
DB	541	TGTGTGGCAACTCGTGGCTTCGTTTGAAGTCCCTATTTCTGGAACCCATTTGATTGT	600	
QY	845	TGTGTCAACTATTGTTTCTCCCTCGTGGCAAGGGGAGGAGGTGCAAGTGGTCTGA	904	
DB	601	TGTGTCAACTATTGTTTCTCCCTCGTGGCAAGGGGAGGAGGTGCAAGTGGTCTGA	660	
QY	905	ACTGATAAAATTTGATGCTGTGGTTCGTTGCCCACTGCTTTCTGGAATTTATGTTCTGG	964	
DB	661	ACTGATAAAATTTGATGCTGTGGTTCGTTGCCCACTGCTTTCTGGAATTTATGTTCTGG	720	
QY	965	AATTTTATTTCTCTGGTTCGTCATATCATCTCCATTAAGGAGATCCAGTTCCTAATGG	1024	
DB	721	AATTTTATTTCTCTGGTTCGTCATATCATCTCCATTAAGGAGATCCAGTTCCTAATGG	780	
QY	1025	TTTGCGAGCTTTGCGCAGTTTCTATGCTCGACAGTTCGAAATTAACCTCTTTTCCATCAT	1084	
DB	781	TTTGCGAGCTTTGCGCAGTTTCTATGCTCGACAGTTCGAAATTAACCTCTTTTCCATCAT	840	
QY	1085	GTATACTGGAG - CACCGTTGCTGGGCTTTGACAAATCTTCCTCTGTGGGGTACCATCTCA	1143	
DB	841	GTATACTGGAGCACCGTTGCTGGGCTTTGACAAATCTTCCTCTGTGGGGTACCATCTCA	900	

RESULT 10

BO669841

LOCUS

DEFINITION

ACCESSION

VERSION

BO669841

AGENCY

BO669841.1

BO669841

BO669841.1

999 bp

linear

EST 15-JUL-2002

5', mRNA sequence.

GI:21780675

Query Match

Best Local Similarity

Matches 892;

26.7%;

99.1%;

Conservative 0;

Score 876.8;

Pred. No. 4.4e-184;

Mismatches 7;

DB 14;

Indels 1;

Gaps 1;

Length 996;

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 999)

NIH-MGC <http://mhc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Place: LNCM2405 row: i column: 06

High quality sequence stop: 618.

BASE COUNT

287 a 221 c 247 g 241 t 3 others

FEATURES

source

1..999

Location/Qualifiers

organism="Homo sapiens"

db_xref="taxon:9606"

clone="IMAGE:6255053"

clone_lib="NIH_MGC_102"

tissue_type="epidermoid carcinoma, cell line"

lab_host="DH10B (phage-resistant)"

note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

Query Match

Best Local Similarity

Matches 954;

Conservative 0;

Mismatches 22;

Indels 7;

Gaps 6;

Length 999;

Score 875.6;

Pred. No. 8.1e-184;

DB 14;

Indels 7;

Gaps 6;

Length 999;

QY

1031

AGCTTTGGCAGTTTCTATGCTCGACAGTTTGAATAAACTCTTTTCCATCATGTATAC

1090

DB

1

AGCTTTGGCAGTTTCTATGCTCGACAGTTTGAATAAACTCTTTTCCATCATGTATAC

60

QY

1091

TGGACACCGTTGCTGGGCTTTGACAACTTCTCTCTGTGGGGTACCATCTCATCTCGT

1150

DB

61

TGGACACCGTTGCTGGGCTTTGACAACTTCTCTCTGTGGGGTACCATCTCATCTCGT

120

QY

1151

GGGATGTCAGTTTCTGTGCCCTTATGCTCTGGTCTTTGTATGTCCAGGATGAAGAG

1210

DB

121

GGGATGTCAGTTTCTGTGCCCTTATGCTCTGGTCTTTGTATGTCCAGGATGAAGAG

180

QY

1211

AAAAATTGAAAGAGAAATAAAGTGTAGTCTTCTGAAAGGCCCTTAATGAAAAAAGAA

1270

DB

181

AAAAATTGAAAGAGAAATAAAGTGTAGTCTTCTGAAAGGCCCTTAATGAAAAAAGAA

240

QY

1271

TAGCTTGAAGAGAACCATGAAGAAACAAAGTCTCTGTGGTGTATTTGAAACCAAGCA

1330

DB</

Db 421 GGACTTTGAAAGAGAAACACAGCATAGATGACACCGTGAATGGTGCAGTGGCTTAA 480

Qy 1511 TGGGAACCTTGTCCAGTTCAGTCAAGCCGTCAGCAACCAAAATAAATCCAGTGGCCACTA 1570

Db 481 TGGGAACCTTGTCCAGTTCAGTCAAGCCGTCAGCAACCAAAATAAATCCAGTGGCCACTA 540

Qy 1571 CCAGTATCACACCGTGCATAGGATTCGGCCCTGTACAAAGAGCTACTCCATAAATTACA 1630

Db 541 CCAGTATCACACCGTGCATAGGATTCGGCCCTGTACAAAGAGCTACTCCATAAATTACA 600

Qy 1631 TCTTGGCAA-GTGGGAGATTCATGGGAGCTCCCGGTGACAAACCCCTTAAGCGCAATA 1689

Db 601 TCTTGGCAAAGGTGGGAGATTCATGGGAGCTCCCGGTGACAAACCCCTTAAGCGCAATA 660

Qy 1690 ATAGCTATCTCTTATACCATGGCAATATGTGGCATGCTCTGGATTCATTCGGTGGCA 1749

Db 661 ATAGCTATCTCTTATACCATGGCAATATGTGGCATGCTCTGGATTCATTCGGTGGCA 720

Qy 1750 AAGAAGGTGAACAGAGGG--CGAAGAAATGGAGAGCTGACATGGCCTAATGCAGACTC 1807

Db 721 AAGAAGGTGAACAGAGGGCGAAGNATGGAGAGCTGACATGGCCTAATGCAGACTC 780

Qy 1808 CAAGAAGCGAATTCGAATGGACAGTACACCACTTACTGCAATGTGTGTCGACCTTCA 1867

Db 781 CAAGAAGCGAATTCGAATGGACAGTACACCACTTACTGCAATGTGTGTCGACCTTCA 840

Qy 1868 CTCAGCATCTGAGATGACATGATGTCAAGGAG-AGATGGGTCTAGG-TGACAGAAA 1925

Db 841 CTCAGCATCTGAGATGACATGATGTCAAGGAGAGATGGGTCTAGGTTGACAGAAA 900

Qy 1926 GGAAGTAATGGCTCTCTAGAAAGATGGTATGA-CCAGGATAAGCCTGAAGTCTCTCTCT 1984

Db 901 GGAAGTAATGGCTCTCTAAAGAAATGGAAATGGAATGACCCAGGATAAAGCTTCTCTCT 960

Qy 1985 -CTTCAGTTCCTGCGAGATCCTT 2006

Db 961 CTTTCCCGGTCTGCGAGACCTT 983

RESULT 11

BQ687887

LOCUS

DEFINITION

AGENCOURT_8341960 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6248984

5', mRNA sequence.

Q687887

Q687887.1 GI:21813203

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 952)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: sgabs-x@mail.nih.gov

Tissue Procurement: ATCC.

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: L1CM2389 row: 1 column: 09

High quality sequence stop: 662.

Location/Qualifiers

1..952

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6248984"

/clone_lib="NIH_MGC_110"

/tissue_type="ductal carcinoma, cell line"

FEATURES

source

/lab_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pOMB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library." 2 others

BASE COUNT 241 a 202 c 240 g 267 t

ORIGIN

Query Match 25.7%; Score 846.6; DB 14; Length 952;

Best Local Similarity 98.2%; Pred. No. 2.2e-177;

Matches 897; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

Qy 723 ATGTACAACTCGACTCAAGGGCTACTGATGCGCGCTCAGTCAGTGTCTATGTTGGTCT 782

Db 1 ATGTACAACTCGACTCAAGGGCTACTGATGCGCGCTCAGTCAGTGTCTATGTTGGTCT 60

Qy 783 GCTGTGGCAACTCGTGGCTTCGTTTGAAGCTCCCTATTCTGGAACCCATTGTATT 842

Db 61 GCTGTGGCAACTCGTGGCTTCGTTTGAAGCTCCCTATTCTGGAACCCATTGTATT 120

Qy 843 GTTGTGCAAACTATTGGTTTCTCCCTCGTGGCAAGGGGCGAGGGTGTCAAGTGGTCT 902

Db 121 GTTGTGCAAACTATTGGTTTCTCCCTCGTGGCAAGGGGCGAGGGTGTCAAGTGGTCT 180

Qy 903 GAACTGATAAAATTTGTGATGTCCTGTCGTCCTTCTGGAATTTATGTCT 962

Db 181 GAACTGATAAAATTTGTGATGTCCTGTCGTCCTTCTGGAATTTATGTCT 240

Qy 963 GGAATTTATTTCTTCTGTCGTCATTCCTTCAAGGAGATCCAGTTCCTTAT 1022

Db 241 GGAATTTATTTCTTCTGTCGTCATTCCTTCAAGGAGATCCAGTTCCTTAT 300

Qy 1023 GGTGTGGCAAGCTTTCGCAAGTTTCTATGCTGCACAGTTGGAATAAACCTCTTTTCCATC 1082

Db 301 GGTGTGGCAAGCTTTCGCAAGTTTCTATGCTGCACAGTTGGAATAAACCTCTTTTCCATC 360

Qy 1083 ATGTATACGTGAGACACCGTGTGCGGCTTTGACAACTTCTCTGTGGGTACCATCTCTC 1142

Db 361 ATGTATACGTGAGACACCGTGTGCGGCTTTGACAACTTCTCTGTGGGTACCATCTCTC 420

Qy 1143 ATCTCGTGGGATGTGCAGTTTCTGTCGCTTATCGTCTGTTGTTGTTGCCAG 1202

Db 421 ATCTCGTGGGATGTGCAGTTTCTGTCGCTTATCGTCTGTTGTTGCCAG 480

Qy 1203 ATGAAGAGAAAAATTTGAACGAGAAAAATAAGTGTAGTCTCTTGAAGCCCTTAAATGGAA 1262

Db 481 ATGAAGAGAAAAATTTGAACGAGAAAAATAAGTGTAGTCTCTTGAAGCCCTTAAATGGAA 540

Qy 1263 AAAAAAGAAATAGCTTGAAGAGACCATGAAGAAAAAAGTGTCTGTTGGTGATATTGAA 1322

Db 541 AAAAAAGAAATAGCTTGAAGAGACCATGAAGAAAAAAGTGTCTGTTGGTGATATTGAA 600

Qy 1323 AACAGCATCTGTTTCTGAGTGGGCTGCGCTGCGCTTCCAGCTGCTGGTGGAG 1382

Db 601 AACAGCATCTGTTTCTGAGTGGGCTGCGCTGCGCTTCCAGCTGCTGGTGGAG 660

Qy 1383 GAGAGAACAGTCTCATTCAAACTTGGAGATTTGGAGGAGCTCCAGAGAGAGAGGCTT 1442

Db 661 GAGAGAACAGTCTCATTCAAACTTGGAGATTTGGAGGAGCTCCAGAGAGAGAGGCTT 720

Qy 1443 CCCAGCTGGACTTGAAGAGAAAAACAGCATATAGTACACCGTGAAT-GGTGCAAGTGA 1501

Db 721 CCCAGCTGGACTTGAAGAGAAAAACAGCATATAGTACACCGTGAATGGTGCAGTGA 780

Qy 1502 GTTGCTTAATGGAACTTGTCCAGTTCAGTCAAGCCGTGACCAACCAATAAATCCAG 1561

Db 781 GTTGCTTAATGGAACTTGTCCAGTTCAGTCAAGCCGTGACCAACCAATAAATCCAG 839

Qy 1562 TGGCCACTACCATGATACACCGT--GCATAAGGATTCGGCTCTGTA-CAAGAGAGCTACT 1618

```
Db      840 TGGNCACTACAGTATCACACCGNTGTCATAGGATTCGGCCCTGTACCAAGAGCTACT 899
Qy      1619 CCATAAATTACAT 1631
Db      900 CCATAAATACAT 912

RESULT 12
BQ423949
LOCUS   BQ423949      896 bp      mRNA      linear      EST 23-MAY-2002
DEFINITION AGENCOURT 7831118 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6151957
5' mRNA sequence.
ACCESSION BQ423949
VERSION   BQ423949.1 GI:21119254
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13489 row: c column: 06
          High quality sequence stop: 729.
          Location/Qualifiers
            1..896
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:6151957"
              /clone_lib="NIH_MGC_67"
              /tissue_type="retinoblastoma"
              /lab_host="DH10B (phage-resistant)"
              /notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
              Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
              Average insert size 1.75 kb. Library constructed by Life
              Technologies."
BASE COUNT 170 a 243 c 220 g 263 t
ORIGIN

Query Match      25.1%; Score 825.8; DB 14; Length 896;
Best Local Similarity 98.7%; Pred. No. 9.4e-173;
Matches 864; Conservative 0; Mismatches 7; Indels 4; Gaps 3;

Qy      11 TGGGTGGTTCTTCTTCTGGCGCGCTGAACCCCGCGGCTCTCTTCTGGAAAGTCTGTGA 70
Db      1 TGGGTGGTTCTTCTTCTGGCGCGCTGAACCCCGCGGCTCTCTTCTGGAAAGTCTGTGA 60

Qy      71 GTCCCGCTGAGCTGTCCCGCGTCCGCGACCCGCGGCGTGTCCCGTGTCCAGCCGC 130
Db      61 GTCCCGCTGAGCTGTCCCGCGTCCGCGACCCGCGGCGTGTCCCGTGTCCAGCCGC 120

Qy      131 TGCCGCTCGATCTCTCTGCTCCCGCTCCGCGCTCCCTTTTCCCTGGATGAACCTTGCCT 190
Db      121 TGCCGCTCGATCTCTCTGCTCCCGCTCCGCGCTCCCTTTTCCCTGGATGAACCTTGCCT 180

Qy      191 CTTTCTCTTCTCGGCCATGAATTCGTCTCCGTGCTTTTAGCCCTCTCTGAGCAAGAA 250
Db      181 CTTTCTCTTCTCGGCCATGAATTCGTCTCCGTGCTTTTAGCCCTCTCTGAGCAAGAA 240

Qy      251 ACCCCAGACAACAGATGCCCATACGCGATATAGCAGTAACCTCCCGAGCTCGGTTCTG 310
Db      241 ACCCCAGACAACAGATGCCCATACGCGATATAGCAGTAACCTCCCGAGCTCGGTTCTG 300
```

```
Qy      311 TGGCGTAGTTTACAGTATTTAAATTTTATATATATATATTTATTTATATAGCATTTTGA 370
Db      301 TGGCGTAGTTTACAGTATTTAAATTTTATATATATATATTTATTTATATAGCATTTTGA 360

Qy      371 TACCTCATATTCGTTTACACATCTTGAAGGGGCTCAGTAGTCTCTTACTAAACACC 430
Db      361 TACCTCATATTCGTTTACACATCTTGAAGGGGCTCAGTAGTCTCTTACTAAACACC 420

Qy      431 ACTACTCCAGAGAAATGGCAACCGTGATTACCACTACTACAGCTGCTACCGCCGCTCTCTGG 490
Db      421 ACTACTCCAGAGAAATGGCAACCGTGATTACCACTACTACAGCTGCTACCGCCGCTCTCTGG 480

Qy      491 TCCTTTGGTGGACTACCTATGAGTGTCTATCCCTGGGCTTCATTATTGCAATTTCTCTTGGC 550
Db      481 TCCTTTGGTGGACTACCTATGAGTGTCTATCCCTGGGCTTCATTATTGCAATTTCTCTTGGC 540

Qy      551 ATTCTCCGTGGGAGCCAAATGATGTAGCAAAATCTTTTGTACAGCTGTGGGCTCAGGTGT 610
Db      541 ATTCTCCGTGGGAGCCAAATGATGTAGCAAAATCTTTTGTACAGCTGTGGGCTCAGGTGT 600

Qy      611 AGTGACCTGAAAGCAAGCCTGCATCCTAGCTAGCATCTTTTGAACACAGTGGGCTCTGTCTT 670
Db      601 AGTGACCTGAAAGCAAGCCTGCATCCTAGCTAGCATCTTTTGAACACAGTGGGCTCTGTCTT 660

Qy      671 ACTGGGGGCAAGTAGTGAGCGAAACCAATCCGGAAGGCTTTGATTGACGTGGAGATGTACAA 730
Db      661 ACTGGGGGCAAGTAGTGAGCGAAACCAATCCGGAAGGCTTTGATTGACGTGGAGATGTACAA 720

Qy      731 CTCGACTCAAGGGCTACTGATGCGCGGCTCAGTCAGTGTCTATGTTGGTCTCTCTGTGTG 790
Db      721 CTCGACTCAAGGGCTACTGATGCGCGGCTCAGTCAGTGTCTATGTTGGTCTCTCTGTGTG 780

Qy      791 GCAACTCGT-GGCTTCGTTTGTGAAGCTCCCTATTTCTGGAACCCATT--GTATTGTGTG 847
Db      781 GCAACTCGTGGGCTTCGTTTGTGAAGCTCCCTATTTCTGGAACCCATTGTATTGTGTG 840

Qy      848 TGCAACTATTGG-TTTCCTCCCTCGTGCAAGGGG 881
Db      841 TGCAACTATTGGTTTTCCTCCCTCGTGCAAGGGG 875

RESULT 13
BQ670753
LOCUS   BQ670753      952 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION AGENCOURT 8192003 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6256942
5' mRNA sequence.
ACCESSION BQ670753
VERSION   BQ670753.1 GI:21781587
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 952)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Ruben Laboratory
          DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM2410 row: g column: 23
          High quality sequence stop: 618.
          Location/Qualifiers
            1..952
              /organism="Homo sapiens"
              /db_xref="taxon:9606"

FEATURES
source
```


Search completed: June 28, 2003, 08:15:50
Job time : 3047.91 secs